economia

Journal of the Latin American and Caribbean Economic Association

Volume 21 Number 2 Spring 2021

Garofalo and Streb on Broken Promises: Regime Announcements and Exchange Rates around Elections

Antón and Rasteletti on Taxing Labor Income in an Economy with High Employment Informality

Alichi, Shibata, and Tanyeri on Fiscal Policy Multipliers in Small States

David, Komatsuzaki, and Pienknagura on Macroeconomic and Socioeconomic Effects of Structural Reforms in Latin America and the Caribbean

O'Leary, Cravo, Sierra, and Justino on Effects of Job Referrals on Labor Market Outcomes in Brazil

Blanco, Cabrera, Carozzi, and Cid on Mandatory Helmet Use and the Severity of Motorcycle Accidents

Volume 21 Number 2

economía

Journal of the Latin American and Caribbean Economic Association

Spring 2021

LATIN AMERICAN AND CARIBBEAN ECONOMIC ASSOCIATION

BROOKINGS INSTITUTION PRESS

Washington, D.C.

Published by
BROOKINGS INSTITUTION PRESS
1775 Massachusetts Avenue, N.W., Washington, D.C. 20036
www.brookings.edu

ISSN 1529-7470 (print)
ISSN 1533-6239 (online)

ISBN 978-0-8157-3969-2 (pbk: alk. paper)
ISBN 978-0-8157-3970-8 (ebook)

For information on subscriptions, standing orders, and individual copies, contact Brookings Institution Press, P.O. Box 465, Hanover, PA 17331-0465. Or call 866-698-0010. Email brookings.tsp@sheridan.com. Visit Brookings online at www.brookings.edu/press.

Brookings periodicals are available online through both OCLC (www.oclc.org) and Project Muse (http://muse.jhu.edu). Archived issues are also available through JSTOR (www.jstor.org).

Volume 21 Number 2

economía

Journal of the Latin American and Caribbean Economic Association

Spring 2021

LATIN AMERICAN AND CARIBBEAN ECONOMIC ASSOCIATION

The Latin American and Caribbean Economic Association (LACEA), or Asociación de Economía de América Latina y el Caribe, is an international association of economists with common research interests in Latin America. It was formed in 1992 to facilitate the exchange of ideas among economists and policymakers. Since 2000, LACEA has been publishing two issues of its own journal, Economía, in the spring and fall of each year. Membership in LACEA is open to all individuals or institutions professionally concerned with the study of Latin American and Caribbean economies. For membership information, please visit the LACEA website at www.lacea.org.

Managing Editor
Elena Zadic

North American Coordinator
Marjorie Pannell

AUTHORS

Ali Mohammad Alichi, *International Monetary Fund*
Arturo Antón, *Bank of Mexico*
Magdalena Blanco, *Universidad de Montevideo*
José María Cabrera, *Universidad de Montevideo*
Felipe Carozzi, *London School of Economics*
Alejandro Cid de Orta, *Universidad de Montevideo*
Túlio Cravo, *African Development Bank*
Antonio David, *International Monetary Fund*
Pablo Garofalo, *New Jersey City University*
Leandro Justino, *World Bank*
Takuji Komatsuzaki, *International Monetary Fund*
Christopher O'Leary, *W. E. Upjohn Institute for Employment Research*
Samuel Pienknagura, *International Monetary Fund*
Alejandro Rasteletti, *Inter-American Development Bank*
Ippei Shibata, *International Monetary Fund*
Ana Cristina Sierra, *World Bank*
Jorge M. Streb, *Universidad del CEMA*
Kadir Tanyeri, *International Monetary Fund*

PABLO GAROFALO
New Jersey City University, United States

JORGE M. STREB
Universidad del CEMA (UCEMA), Argentina

Broken Promises: Regime Announcements and Exchange Rates around Elections

ABSTRACT We study exchange rate dynamics around government changes conditional on the exchange rate regime, which we identify by combining the IMF de jure and the Reinhart and Rogoff de facto exchange rate regime classifications. This allows distinguishing whether the official exchange rate regime announcements match actual policy or are inconsistent with it. Using monthly data from Latin American democracies, we do not find significant exchange rate depreciations before the change of government in any of the regimes we identify. However, we do detect a gradual real exchange rate overvaluation when the de jure regime is fixed but the de facto policy is flexible, which is abruptly corrected after the change of government; this pattern of real exchange rate misalignments when the announcement does not match actual behavior is linked to incumbents that postpone devaluations until the successor steps in. This pattern of broken promises is typical until 1999, but it becomes exceptional thereafter.

JEL Codes: D72, D78, E00

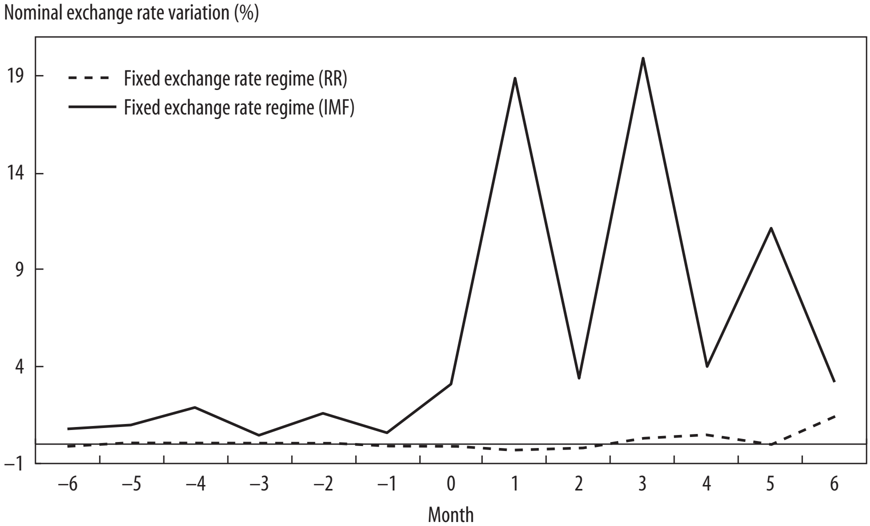
Keywords: Exchange rates, exchange rate misalignment, exchange rate regimes, electoral cycles

Reneging on exchange rate regime announcements occurs quite often. We track exchange rate regime announcements with the de jure exchange rate classification maintained by the International Monetary Fund (IMF), which reports what countries claim to be doing.[1] The IMF de jure classification has been criticized for representing words, not deeds (Levy-Yeyati and Sturzenegger, 2005; Reinhart and Rogoff, 2004). Among the de facto classifications proposed, Reinhart and Rogoff (2004) reclassify exchange rate arrangements by developing an algorithm based on the observed behavior of

ACKNOWLEDGMENTS The editor, Veronica Rappoport, and two reviewers helped us to clarify the conceptual framework and sharpen the empirical results. We also appreciate the comments of Juan Pereyra and Jennifer Hoover.

1. Exchange rate regime announcements are here distinguished from firmer monetary commitments like dollarization, in which the country relinquishes an independent currency, for example, Panama since 1904, Ecuador since 2000, and El Salvador since 2001. A regime announcement continues to hold until further notice.

FIGURE 1. Exchange Rate Devaluations around Government Changes



Note: Average exchange rate variation during nineteen and twenty-one complete episodes computed with the RR and IMF fixed exchange rate classifications, respectively, in twenty-one Latin American countries (Argentina, Barbados, Bolivia, Brazil, Chile, Colombia, Costa Rica, Dominican Republic, Ecuador, El Salvador, Guatemala, Guyana, Honduras, Jamaica, Mexico, Nicaragua, Paraguay, Peru, Trinidad and Tobago, Uruguay, and Venezuela) in the 1980–99 period. Both classifications are invariant throughout the twelve-month window. Month 0 is the government change month. Dollarization episodes are excluded.

exchange rates, while parallel exchange rates are used if multiple markets are present. While Reinhart and Rogoff (2004, p. 1) claim that the IMF exchange rate classification is "a little better than random," we have reasons to suspect otherwise.

Using the IMF and Reinhart and Rogoff (RR) classifications, figure 1 shows nominal exchange rate variations around government changes (when an incumbent's term ends and a new administration is inaugurated) in Latin American countries, conditional on a fixed exchange rate regime within the whole window. Devaluations are similar under both classifications up to the month of government change, but they increase considerably thereafter under the IMF de jure classification. This suggests that some exchange rate pegs are sustained in the prelude to elections and government changes, but not afterward.

Calvo and Reinhart (2002) show that many Latin American countries that claim to be floating are not doing so, a phenomenon known as fear of floating. This occurs, for instance, when a country classified as floating is in reality pegging its exchange rate to, say, the U.S. dollar. Conversely, Alesina and Wagner (2006) show that some countries often break commitments to pegging and end up floating more than they announce, a phenomenon they call fear of pegging. We analyze these mismatches around elections. Inspired by Alesina and Wagner (2006), we combine the de jure and de facto classifications to distinguish between "keeping" and "breaking" promises, that is, between regime announcements that are consistent with observed market-based exchange rate behavior and those that are not. As Genberg and Swoboda (2005) note, both announcements and actions may provide useful information about exchange rate policy. We thus explore the behavior of exchange rates conditional on the regimes that we identify based on the consistency of the de jure and de facto classifications. To the best of our knowledge, nobody has analyzed this issue before. While there is ample evidence on the delay of exchange rate adjustments when elections are coming up (for example, Cermeño, Grier, and Grier, 2010; Edwards, 1994; Stein and Streb, 2004; Stein, Streb, and Ghezzi, 2005), these earlier studies may suffer from downward bias because they do not control for either exchange rate regimes or the consistency between the de jure and de facto classifications: their results are a weighted average of devaluations in inconsistent de jure fixed exchange rate regimes, where we find that all the variability is concentrated, and all the other regimes, where no pattern is found. We henceforth focus on these inconsistent de jure fixed regimes, those exhibiting the so-called fear of pegging.[2] Observationally, de jure fixed regimes that are inconsistent with the de facto flexible behavior share identifiable underlying characteristics, namely, dual or multiple markets and high inflation before elections. These inconsistent fixed regimes, for short, always involve broken promises. In contrast, fear of floating need not imply broken promises: Genberg and Swoboda (2005) point out that a country that may seem to be pegging its currency to another country's might simply be following a similar monetary policy, so it is not breaking any commitment.

We first study the determinants of exchange rate regimes around elections using ordered logit models for both the IMF de jure and RR de facto regime classifications. As found by Klein and Marion (1997) and Gavin and Perotti (1997), there is no evidence that de jure regimes change before the government turnover, but the probability of abandoning a fixed exchange rate regime increases after the new administration is inaugurated. Additionally, we find that the probability of a de facto flexible regime increases before government

2. Inconsistent de jure fixed regimes are a slight modification of what Alesina and Wagner call fear of pegging. We develop the rationale for this modification later in the paper.

changes. We then rely on a multinomial logit model, which is widely used for unordered categories, to study the consistency between de jure and de facto regimes. This is something novel in the literature on exchange rate regimes. After government changes, we detect that the probability of inconsistent fixed regimes, which are de facto flexible, decreases in relation to the probability of consistent flexible regimes, with are both de jure and de facto flexible. In other words, though the market regime behavior already involves some degree of float, the authorities announce it after the inauguration of a new government, not before.

Second, we study the dynamics of the real exchange rate around the government change month, conditional on whether the de jure regime matches the de facto regime before the month of the election, by using a dynamic distributed lag model and a difference-in-differences strategy. We find that exchange rate behavior during consistent and inconsistent de jure fixed regimes is not statistically different until the month of the government turnover, but it differs significantly in the first quarter after that. Hence, although inconsistent fixed regimes might tend to be episodes of "poor macroeconomic performance and inability to maintain monetary and fiscal stability" (Alesina and Wagner, 2006, p. 774), official exchange rates are sustained until the government change date.

Third, the paper contributes to the literature on real exchange rate appreciations and their reversions. Goldfajn and Valdés (1999) show that real exchange rate appreciations are usually reverted by nominal exchange rate devaluations rather than by smooth inflation differentials. This nominal adjustment through sharp exchange rate devaluations causes overvaluation to last longer during the buildup stage than during the reversion stage. In our sample of Latin American countries, the overvaluation of the real exchange rate occurs only for inconsistent fixed regimes. Such overvaluation begins ten months before the government change date and lasts until two months after the government turnover (about one year of overvaluation), with a peak of 37 percent in the government change month. Reversion starts abruptly the next month and is completed in three months. This corroborates the findings by Goldfajn and Valdés (1999) on the asymmetry between the buildup and reversion stages due to sudden nominal exchange rate adjustments. While they do not characterize and describe the context in which these appreciation episodes take place, we identify one particular context where they occur: inconsistent fixed regimes before government changes.

Finally, we compare the 1980–99 period, to which the findings described above apply, with the 2000–16 period. Although the IMF changed to a de facto classification after 1999, we can use the fact that dual/multiple regime practices are an underlying characteristic of inconsistent de jure fixed regimes to study the recent period. We find that dual/multiple regimes are almost nonexistent in the 2000–16 period, but we provide a case study of the only election where the regime is likely to be classified as inconsistent fixed, namely, the Argentine general election of 2015.

The rest of the paper is organized as follows. The next section reviews the exchange rate classification literature. We then explain our methodology for identifying consistent and inconsistent exchange rate regime announcements. Subsequently we present the econometric models and results, analyze the appreciation of the real exchange rate and its reversion, and compare the 1980–99 period with the most recent period by relying on the underlying characteristics of inconsistent fixed regimes. The final section concludes.

Exchange Rate Regime Classifications

The IMF developed a traditional exchange rate regime classification, which it has published since 1950.[3] Until 1999, it asked "country members to self-declare their arrangement as belonging to one of four categories" (Alesina and Wagner, 2006, p. 775): float, managed, crawl, or fixed. If a country announced the adoption of a floating regime in a specific year, "the IMF classified this country-year as floating even if in practice this country pegged its currency to, say, the U.S. dollar" (Alesina and Wagner, 2006, p. 775). There are many reasons to seek other approaches to classifying exchange rate regimes. For instance, empirical work on the costs and benefits of alternative exchange rate arrangements can be misleading when actual behavior deviates significantly from the announced behavior; as Reinhart and Rogoff (2004) point out, Baxter and Stockman (1989) find that there are no significant differences in business cycles across exchange arrangements.

Reinhart and Rogoff (2004) provide a "natural classification" of exchange rate regimes that relies on a broad variety of descriptive statistics to group episodes into a grid of regimes based on market-determined exchange rate behavior. They provide detailed analyses to posit the importance of market-determined exchange rates as the best indicator of the underlying monetary policy. They first do so by showing that the market exchange rate consistently

3. *Annual Report on Exchange Arrangements and Exchange Restrictions*.

anticipates devaluation of the official rate, and not vice versa. Second, they find that the market-determined exchange rate keeps up with inflation, while the official rate sometimes does not. Additionally, they remark that "it is not unusual for dual or parallel markets (legal or otherwise) to account for the lion's share of transactions with the official rate being little more than symbolic" (Reinhart and Rogoff, 2004, p. 10).

To create the natural classification, they first check whether there is a unified rate instead of dual or parallel (black) markets. If there is a dual or parallel market, given the relevance of the market-determined rate explained above, they classify the regime as de facto using the market-determined exchange rate. If there is no parallel market, they examine summary statistics to verify the official de jure arrangement, if any, going forward from the date of the announcement. If the regime is verified, it is then classified as de jure. If the de jure regime fails verification, they seek a de facto statistical classification based on the behavior of the exchange rate if inflation is below 40 percent. When annual inflation is above 40 percent, the exchange rate is classified as "freely falling." A similar statistical classification is conducted when there is no preannounced path for the exchange rate. In all, they establish fourteen categories in what they call their fine grid, which they collapse into five categories for their coarse grid. We use the latter in our analysis.

Levy-Yeyati and Sturzenegger (2005) also provide a de facto classification of exchange rate regimes. Besides exchange rates, their algorithm uses base money and international reserves. While both classifications have their merits, the RR classification suits our analysis better because it provides a monthly classification that allows us to observe switching regimes, if any, around elections and government change dates, which is important to determine the endogeneity of the regime. Moreover, Levy-Yeyati and Sturzenegger (2005) use the official exchange rate in their de facto algorithm, rather than market rates. Like Alesina and Wagner (2006, p. 797), we are interested in how de facto behavior deviates from announced official policies, so this also points to the RR classification.

Consistency of De Jure and De Facto Exchange Rate Regimes

To identify consistent and inconsistent de jure regimes (that is, whether the official announced regime matches the actual policy), we follow an approach similar to Alesina and Wagner (2006), who quantify broken promises, which we call inconsistencies, as the difference between the coarse RR and IMF

FIGURE 2. Classification of De Jure Exchange Rate Regimes by Alesina and Wagner (2006)

RR de facto classification (actual policy)

Float	4, 1 (+3)	4, 2 (+2)	4, 3 (+1)	4, 4 (0)
Managed	3, 1 (+2)	3, 2 (+1)	3, 3 (0)	3, 4 (–1)
Crawl	2, 1 (+1)	2, 2 (0)	2, 3 (–1)	2, 4 (–2)
Fixed	1, 1 (0)	1, 2 (–1)	1, 3 (–2)	1, 4 (–3)
	Fixed	Crawl	Managed	Float

IMF de jure classification (announcement)

Source: Alesina and Wagner (2006).

Note: Each cell contains three numbers: *X*, *Y* (*Z*). *X* represents the RR classification and *Y* the IMF classification (where 4 = float, 3 = managed, 2 = crawl, and 1 = fixed), and $Z = X - Y$. Dark gray: Fear of floating, with more management than announced ($Z < 0$). Light gray: Fear of pegging, with more floating than announced ($Z > 0$).

classifications. They assign a value from one to four to identify the regime (fixed: 1; crawl: 2; managed: 3; float: 4) and then subtract the de jure value from the de facto value. For example, if the RR natural classification of the regime is a float (with a value of 4) while the IMF de jure classification is managed (value of 3), then the difference (denoted *Z*) is positive and called fear of pegging. A negative value for *Z*, in turn, indicates fear of floating, while a value of zero represents a consistent regime. Figure 2 shows all the possible combinations.

This classification does not control for the intensity of the differences between the RR and IMF classifications. It applies equally to $Z = -3$ and $Z = -1$, without distinguishing between strong and weak fear of floating (an analogous observation holds for $Z > 0$ regarding the different intensities of fear of pegging). This issue is the starting point for our regime classification below. Our main innovation lies in dividing consistent de jure regimes into fixed (fixed or crawl) and flexible (managed or float). We create the categories using a two-dimensional classification system: fixed versus flexible and consistent versus inconsistent. Our approach is depicted in figure 3, which presents four categories of de jure regimes: (1) consistent fixed (intermediate gray), (2) consistent flexible (unshaded), (3) inconsistent fixed (strong fear

FIGURE 3. Alternative Classification of De Jure Exchange Rate Regimes

RR de facto classification (actual policy)

	Fixed	Crawl	Managed	Float
Float	4, 1 (+3)	4, 2 (+2)	4, 3 (+1)	4, 4 (0)
Managed	3, 1 (+2)	3, 2 (+1)	3, 3 (0)	3, 4 (−1)
Crawl	2, 1 (+1)	2, 2 (0)	2, 3 (−1)	2, 4 (−2)
Fixed	1, 1 (0)	1, 2 (−1)	1, 3 (−2)	1, 4 (−3)

IMF de jure classification (announcement)

Note: Each cell contains three numbers: *X*, *Y* (*Z*). *X* represents the RR classification and *Y* the IMF classification (where 4 = float, 3 = managed, 2 = crawl, and 1 = fixed), and $Z = X - Y$. Dark gray: inconsistent flexible ($Z < -1$). Light gray: inconsistent fixed ($Z > 1$). Intermediate gray: consistent fixed. Unshaded: consistent flexible.

of pegging; light gray), and (4) inconsistent flexible (strong fear of floating; dark gray). The consistent de jure regime categories correspond to $Z = 0$ and $Z = \pm 1$ when there is either a match between the actual policy and the de jure regime ($Z = 0$) or a weak departure ($Z = \pm 1$). This is how we differentiate the intensity of the episodes in our analysis; that is, $Z \leq \text{abs}(1)$ belongs to consistent de jure regimes, while $Z > \text{abs}(1)$ belongs to inconsistent ones.

Data, Econometric Specifications, and Results

Our main focus is real exchange rate dynamics around government change dates conditional on the consistency of the de jure exchange rate regimes. We first study the determinants of the exchange rate regime policies and the extent to which they are sensitive to the electoral window. This is an important issue to address since regime types are used as controls in the study of exchange rate dynamics. Therefore, netting out covariates, we would like to see how endogenous regimes are around government changes, if at all. We can only carry out these econometric analyses for the 1980–99 period because the IMF abandoned its de jure classification after that.

We collected monthly data on exchange rates and inflation from twenty-one Latin American countries from the IMF International Financial Statistics (IFS) database over the 1980–99 period. The countries are Argentina, Bolivia, Brazil, Chile, Colombia, Costa Rica, Dominican Republic, Ecuador, El Salvador, Guatemala, Guyana, Honduras, Jamaica, Mexico, Nicaragua, Panama, Paraguay, Peru, Trinidad and Tobago, Uruguay, and Venezuela.[4] We constructed the series of the multilateral real exchange rate, which is a trade-weighted average of bilateral real exchange rates. We follow Goldfajn and Valdés (1999) in using only trading partners above 4 percent of overall trade. Also as in Goldfajn and Valdés (1999), we fixed the trade weights using trade flows of an intermediate year (1995 in our case) from the United Nations *International Trade Statistics Yearbook*.[5] Monthly observations of the RR natural exchange rate regime classification are from Ethan Ilzetzki's website, an updated version of the original data from Carmen M. Reinhart's website.[6] The traditional IMF annual exchange rate regime classification comes from the IMF *Annual Report on Exchange Arrangements and Exchange Restrictions (AREAER)*.[7] We conducted a country-by-country study to transform the IMF annual classification into monthly series by reviewing all AREAER manuals from 1980 to 1999 (see details, methodology, and sources in online appendix A).[8]

Duration of Exchange Rate Regimes

We compare regime duration inside and outside the electoral window, because exchange rate estimations controlling for regime at election time may be biased

4. Chile, El Salvador, Guyana, Jamaica, Paraguay, and Trinidad and Tobago are dropped from the sample when the full set of covariates is used owing to missing observations in control variables for these countries. We thus work with two samples: a reduced sample that excludes these countries and an extended sample that includes them. Results for the reduced sample are very similar with and without covariates. Results for the extended sample are only available without covariates. While results without controls are somewhat smaller in magnitude for the extended sample, they are significant in both cases.

5. Identical qualitative results were found using only the bilateral real exchange rate with the United States. This may be because the United States is the main trading partner for almost all Latin American countries. We therefore conclude that our results should not be sensitive to the year of weights used. These alternative results are available on request.

6. Ethan Ilzetzki: www.ilzetzki.com/irr-data; Carmen M. Reinhart: www.carmenreinhart.com. See Ilzetzki, Reinhart, and Rogoff (2019).

7. Available online at 0-www-elibrary-imf-org.library.svsu.edu/subject/012?t=F&type_0=book&type_1=journalissue.

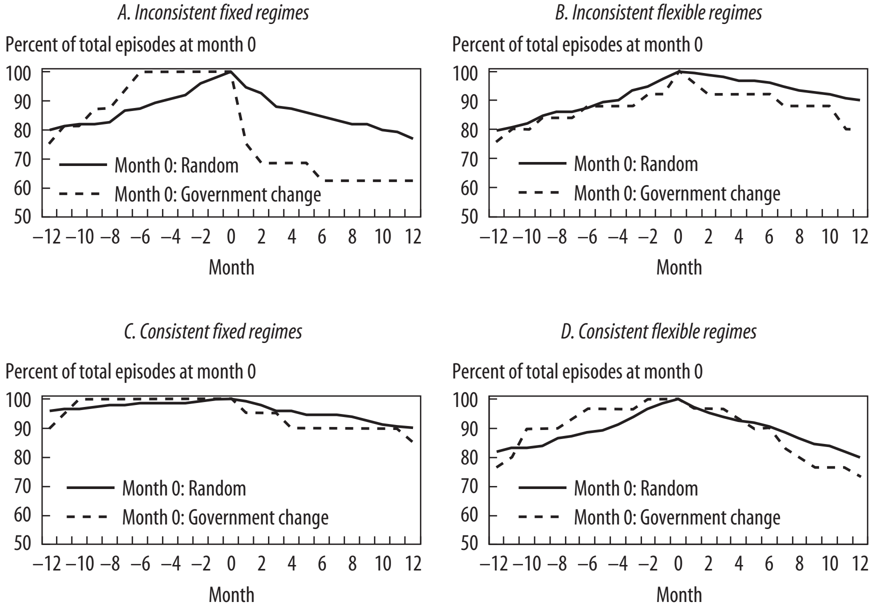
8. Supplementary material for this paper is available online at http://economia.lacea.org/contents.htm.

if regime duration is sensitive to the electoral window. We proceed to report summary statistics for our four-category regime classification: (1) inconsistent fixed (that is, de jure fixed in the IMF classification and de facto flexible in the RR classification); (2) consistent fixed (both de jure and de facto fixed); (3) inconsistent flexible (de jure flexible and de facto fixed); and (4) consistent flexible (both de jure and de facto flexible).

To compare regime duration inside and outside the electoral window, we generate, with a uniform distribution, a random number between one and 240 in order to select a month in the 1980–99 period for a given country. We then observe the regime classification for that month and construct a twenty-four-month window around the observation (twelve months on either side) to identify the duration of that regime. For example, an episode might change at month –5 (that is, the classification at month 0 started five months earlier) or at month +5 (that is, the classification at month 0 ended five months later). We repeat this randomization fifty times for each country and then, for each month, calculate the percentage of episodes in which the regime continues to equal that of month 0. This randomization process allows us to have an idea of the duration of regimes, independently of the covariates. We conduct exactly the same exercise around government change months. The percentage of episodes in each category is very similar for the two windows: 15.4 percent (17.6 percent) are inconsistent fixed regimes for the random (government change) month 0; 28.3 percent (27.5 percent) are inconsistent flexible regimes; 22.9 percent (22.0 percent) are consistent fixed regimes; and 33.5 percent (33.0 percent) are consistent flexible regimes. This similarity suggests that the distribution of regimes is not sensitive to the electoral window. Results are displayed in figure 4.

When looking at the duration of the regimes in figure 4, about 95 percent of the episodes of consistent fixed regimes (panel C) were already in that category eleven months before month 0 for both windows, and about 90 percent of the episodes continue to belong to this category eleven months later. Distributions are also similar for inconsistent flexible (panel B) and consistent flexible regimes (panel D). The exception is inconsistent fixed regimes (panel A). Cases in this category appear to have a longer duration before month 0 for the government change window: for instance, 100 percent of the government change episodes in this category had already started by month –6, versus only 87 percent for the random data window. Although this sounds problematic, in the next subsection we show that, after netting out covariates, the probability of inconsistent fixed regimes does not increase when government change approaches. In contrast, the abrupt decrease after government changes

FIGURE 4. Duration of Regime Classifications around Random and Government Change Months



Notes: Number of regime episodes at government change month 0: inconsistent fixed, 16; inconsistent flexible, 25; consistent flexible, 30; consistent fixed, 20. Number of regime episodes at random month 0: inconsistent fixed, 169; inconsistent flexible, 311; consistent flexible, 368; consistent fixed, 252.

anticipates a finding below: some incumbents switch categories from inconsistent fixed to consistent flexible.

Determinants of Exchange Rate Regimes

We use two estimation methods in this subsection. We first study the IMF de jure and the RR de facto regime classification separately for the period 1980–99 to compare with the findings of previous studies. Since both classifications represent clear ranks with a meaningful order (that is, fixed, crawl, managed, and float), ordered logit models are the appropriate option. We then study the determinants of our novel regime classification, that is, de jure fixed and flexible regimes that are either consistent or inconsistent, as shown in figure 4. Since it is hard to construct a meaningful order, the most suitable option is to adopt the multinomial logit model, where the probability of a category is computed in relation to a selected base category. We use inconsistent de jure fixed regimes as the base category. For both estimation methods, we use the same set of covariates to identify regime determinants, based on the following set of conditional probabilities:

$$\text{(1)} \qquad \mathrm{P}\left[Y_{it} = y \middle| \mathbf{X}_{\mathbf{it}}, \mathbf{GovCh}\left(\mathbf{q}^{-}\right)_{\mathbf{it}}^{\mathbf{q}}, \mathbf{GovCh}\left(\mathbf{q}^{+}\right)_{\mathbf{it}}^{\mathbf{q}}\right],$$

where i and t stand for country and month, respectively. For the ordered logit model, where the categories present a clear rank order, the dependent variable is either the de jure or the de facto exchange rate regime, and y takes a value of 1, 2, 3, or 4 if the regime is fixed, crawl, managed, or float, respectively. For the multinomial logit model, where the categories are not ruled by any apparent rank order, the dependent variable y takes a value of 1, 2, 3, or 4 if the de jure regime is consistent fixed, inconsistent flexible, inconsistent fixed, or consistent flexible, respectively. Government change is represented by two matrices of four dummy variables each, accounting for the year before and after the change, which occurs in month 0. Although the data are monthly, we define the dummy variables by quarters (the superscript q stands for quarter). Thus,

$$\mathbf{GovCh}\left(\mathbf{q}^{-}\right)_{\mathbf{it}}^{\mathbf{q}} = \left[GovCh(-3)_{it}^{q}\ GovCh(-2)_{it}^{q}\ GovCh(-1)_{it}^{q}\ GovCh(0)_{it}^{q}\right],$$

where $GovCh(0)^q$ takes a value of one in the months 0 to 2 before the government change month, $GovCh(-1)^q$ takes a value of one in the months 3 to 5 before the government change month, and so on for $GovCh(-2)^q$ and $GovCh(-3)^q$. Analogously,

$$\mathbf{GovCh}\left(\mathbf{q}^{+}\right)_{\mathbf{it}}^{\mathbf{q}} = \left[GovCh(+1)_{it}^{q}\ GovCh(+2)_{it}^{q}\ GovCh(+3)_{it}^{q}\ GovCh(+4)_{it}^{q}\right]$$

is constructed for the twelve months following the change in government using four quarterly dummy variables.

X is a matrix composed of seven time-varying controls: (1) *Portfolio*: the sum of the absolute value of inward and outward flows of portfolio investment and financial derivatives as a percentage of GDP, from the IMF International Financial Statistics (IFS) database. Levy-Yeyati, Sturzenegger, and Reggio (2010) use this variable as a proxy for capital mobility. Given the impossible trinity, policymakers must give up on either monetary policy or exchange rate policy in environments with high capital mobility, which makes intermediate regimes less viable. Alternatively, under a currency mismatch argument, we should expect more commitments to pegging. (2) *Foreign.Liab.pc*:

foreign liabilities per capita, from the IMF IFS database. Countries with substantial foreign liabilities may be prone to fix their currency, since sharp nominal depreciation affects the solvency of the nontradable sector's balance sheets. Alesina and Wagner (2006) and Levy-Yeyati, Sturzenegger, and Reggio (2010) use foreign liabilities over monetary aggregates instead. However, the problem with this variable is that for Latin American countries, money demand was extremely unstable in the 1980s and early 1990s owing to high inflation. In crisis episodes with high inflation, money demand falls while the monetary authority lets the exchange rate float, creating a positive relation between foreign liabilities and flexible regimes, totally opposite to the currency mismatch hypothesis.[9] (3) *Size*: real GDP in dollars, from the IMF IFS database. As noted by Levy-Yeyati, Sturzenegger, and Reggio (2010), smallness favors a more stable exchange rate because small economies are more likely to trade internationally than economies with a large domestic market and because it limits the scope for the use of a national unit of account. (4) *ToT:* terms of trade. When the terms of trade are high, Latin American countries tend to fix their exchange rates as a device for accumulating international reserves in their central banks, probably to insure against sudden stops (Jeanne, 2007; Jeanne and Rancière, 2011). (5) *U.S.Interest*: the U.S. interest rate in real terms, from the IMF IFS database. Calvo, Leiderman, and Reinhart (1993) and Fernández-Arias and Montiel (1996) find that the U.S. interest rate is a determinant of capital inflows in Latin America.[10] When the U.S. interest rate increases, capital outflows may be stopped by letting the exchange rate float. This effect should be exacerbated when economies keep more open capital accounts. (6) *Openness*: exports plus imports over GDP, from the IMF IFS database. The decision to peg could be correlated with trade openness since highly open economies are in favor of a more stable exchange rate, as Levy-Yeyati, Sturzenegger, and Reggio (2010) note. (7) *Default*: a dummy variable that takes a value of one if the country has defaulted on its external debt and zero otherwise, from Carmen M. Reinhart's website. This variable is used to control for the fact that economies characterized by high macroeconomic instability cannot sustain their currency, so they let their currency float or, more precisely, freely fall.

9. We indeed find a significant positive coefficient when we use foreign liabilities over money, so the probability of a flexible regime increases when foreign liabilities to money increase. In contrast, there is a negative coefficient with our transformation of foreign liabilities normalized by population. The latter is consistent with the currency mismatch hypothesis as found in Levy-Yeyati, Sturzenegger, and Reggio (2010) for their regime classification. Results are shown below.

10. When the U.S. Treasury bill rate is used instead, results are qualitatively the same.

Among the seven controls described above, five are available only at annual frequencies. These are *Portfolio*, *Foreign.Liab.pc*, *Size*, *ToT*, and *Default*. For the first four, we use the log differential method to construct within-year imputation with constant monthly percentage change within each year. *Default* is left at its annual frequency insofar as it is a dummy variable. The remaining two, *U.S.Interest* and *Openness,* are available at monthly frequencies, so interpolation is not necessary. Given the possibility of reverse causality, we decided to use one-month lagged values of the variables available at monthly frequency. For the variables available at annual frequency that were interpolated using log differences, we use twelve-month lagged values instead. All variables are expressed in natural logs except *Default* and dummy variables for government change.

The estimations of equation 1 under the IMF de jure and the RR de facto exchange rate regimes for ordered logit models are shown in table 1, together with the results for the multinomial logit model. Results of the ordered and multinomial logits have different interpretations. For the former, a positive (negative) estimator indicates that the probability of a more flexible (fixed) regime increases if the corresponding covariate shows a marginal increase, but the estimator does not predict at first sight what happens with the probabilities of the middle categories.[11] For the latter, each coefficient is understood as the increase in the probability of category $j = 1, 2, 4$ in relation to the base category (3) for a marginal increase of the independent variable, if its coefficient is positive.

Here we first focus on the econometric results in table 1 for each of the covariates and relate them to the well-known literature on the de jure and de facto regime determinants.[12] Our innovation in relation to the literature is the novel regime classification, which identifies consistent and inconsistent de jure regimes (results are displayed in columns 3–5 of table 1). We then focus on the issue of endogeneity of regimes around the government change date.

For the RR classification in column 2, the probability of observing de facto fixed regimes tends to increase as the de facto capital account openness increases (that is, *Portfolio* = –0.045** in column 2). This is consistent with the currency mismatch hypothesis of Levy-Yeyati, Sturzenegger, and Reggio (2010). At the same time, de jure flexible regimes also tend to increase as the de facto

11. See online appendix B for the full set of marginal effects for each of the categories using mean values of covariates for both the ordered and multinomial logit models.

12. The connection between regression coefficients and changes in probabilities is detailed in online appendix B.

TABLE 1. Determinants of Exchange Rate Regimes: Ordered and Multinomial Logit Models

	Ordered logit		*Multinomial logit relative to inconsistent de jure fixed*		
Explanatory variable	*IMF (de jure) (1)*	*RR (de facto) (2)*	*Consistent de jure fixed (3)*	*Inconsistent de jure flexible (4)*	*Consistent de jure flexible (5)*
ln $Portfolio_{t-12}$	0.046**	−0.045**	0.166***	0.168***	0.104***
	(0.021)	(0.019)	(0.033)	(0.043)	(0.030)
ln $Foreign.Liab.pc_{t-12}$	0.012	−0.049***	0.265***	0.201***	0.297***
	(0.009)	(0.009)	(0.016)	(0.022)	(0.018)
ln $Size_{t-12}$	−0.006	−0.041	1.403***	1.172***	1.576***
	(0.041)	(0.040)	(0.105)	(0.110)	(0.104)
ln ToT_{t-12}	−0.656	−2.564***	4.141***	10.357***	3.193**
	(0.485)	(0.438)	(1.383)	(1.379)	(1.310)
ln $U.S.Interest_{t-1}$	−2.333***	−0.017	−0.897***	−4.550***	−2.517***
	(0.135)	(0.124)	(0.219)	(0.248)	(0.237)
ln $Openness_{t-1}$	0.464***	−0.297***	0.883***	0.665***	1.331***
	(0.109)	(0.105)	(0.222)	(0.232)	(0.227)
$Default_{t-12}$	0.588***	1.573***	−0.660***	−0.493*	1.604***
	(0.100)	(0.099)	(0.227)	(0.273)	(0.221)
$GovCh(-3)^q$	0.048	0.297*	−0.025	0.040	0.481
	(0.166)	(0.160)	(0.383)	(0.414)	(0.395)
$GovCh(-2)^q$	0.125	0.434***	−0.434	−0.176	0.288
	(0.169)	(0.163)	(0.341)	(0.379)	(0.351)
$GovCh(-1)^q$	0.163	0.449***	−0.503	−0.326	0.228
	(0.170)	(0.163)	(0.321)	(0.372)	(0.327)
$GovCh(0)^q$	0.201	0.464***	−0.530	−0.227	0.213
	(0.172)	(0.166)	(0.325)	(0.366)	(0.326)
$GovCh(+1)^q$	0.595***	0.630***	0.183	0.229	0.985**
	(0.172)	(0.167)	(0.397)	(0.431)	(0.387)
$GovCh(+2)^q$	0.483***	0.496***	0.246	0.341	0.838**
	(0.169)	(0.162)	(0.399)	(0.421)	(0.392)
$GovCh(+3)^q$	0.346**	0.322**	0.249	0.020	0.411
	(0.169)	(0.158)	(0.378)	(0.414)	(0.386)
$GovCh(+4)^q$	0.396**	0.255	0.106	0.080	0.080
	(0.172)	(0.160)	(0.375)	(0.410)	(0.376)
No. observations	2,557	2,592	2,662	2,662	2,662

* $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Notes: Columns 1 and 2 show the results of the estimation of equation 1 with ordered logit models, where the dependent variable equals 1, 2, 3, or 4 if the regime is fixed, crawl, managed, or float, respectively. Columns 3, 4, and 5 show the results of the estimation of equation 1 with multinomial logit, where the dependent variable equals 1, 2, 3, or 4 if the de jure regime is consistent fixed, inconsistent flexible, inconsistent fixed, or consistent flexible, respectively; the results are relative to the inconsistent fixed category. Reduced sample includes Argentina, Bolivia, Brazil, Colombia, Costa Rica, Dominican Republic, Ecuador, El Salvador, Guatemala, Honduras, Mexico, Nicaragua, Peru, Uruguay, and Venezuela over the 1980–99 period. Nondemocratic episodes are excluded, based on the Polity IV Project. Dollarization episodes are also excluded. Robust standard errors are in parentheses.

capital account openness increases (that is, *Portfolio* = 0.046** in column 1). Taken together, these results may be suggestive of an increase in inconsistent flexible regimes. In columns 3 to 5, we find that the facto fixed regimes, including both the consistent fixed and inconsistent flexible versions, are more likely relative to inconsistent fixed regimes (that is, 0.166*** and 0.168***, respectively). Moreover, inconsistent fixed regimes are less likely since all three coefficients are positive (that is, all regimes are more likely in relation to the base category).

The coefficient for foreign liabilities per capita (*Foreign.Liab.pc*) is close to zero and insignificant for the IMF de jure classification in column 1, but it is significantly negative in the de facto classification in column 2 (that is, *Foreign.Liab.pc* = –0.049***), which is consistent with the currency mismatch hypothesis. The multinomial logit model corroborates this finding: both types of fixed regimes—consistent fixed (column 3) and inconsistent flexible (column 4)—are more likely relative to inconsistent fixed regimes (that is, 0.265*** and 0.201***, respectively). Consistent flexible regimes are also more likely relative to inconsistent fixed regimes (that is, 0.297*** in column 5). This suggests that inconsistent fixed regimes do not go hand in hand with liability dollarization, mainly because those are episodes of high macroeconomic instability. *Size* is insignificant in both columns 1 and 2, while we would have expected it to be positive at least for the de facto classification. However, in line with Levy-Yeyati, Sturzenegger, and Reggio (2010), the multinomial logit model finds that the consistent flexible category increases its likelihood the most, insofar as its estimator is the greatest of the three. This indicates that flexible regimes are indeed more likely in bigger countries, while inconsistent fixed regimes become less likely because the three estimators are positive. *ToT* has the predicted negative sign in the de facto classification of column 2, while in the de jure classification it is close to zero and insignificant. In columns 3–5, the two de facto fixed regimes—that is, consistent fixed (column 3) and inconsistent flexible (column 4)—become more likely when terms of trade increase. This is consistent with the strategy of pegging the exchange rate to acquire international reserves as, probably, an insurance device, as found in Jeanne (2007) and Jeanne and Rancière (2011). In addition, since all three estimators are positive, it indicates that an inconsistent fixed regime becomes less likely. This is probably because the increase in terms of trade tends to create a trade balance surplus that increases the supply of foreign currency, which may alleviate exchange rate pressures during high macroeconomic instability.

U.S.Interest is close to zero and insignificant in the de facto regime (column 2). For the de jure regime (column 1), the likelihood of a peg increases strongly as *U.S.Interest* increases, since the estimator is significant and negative. Altogether, this evidence might indicate the increase of de jure fixed regimes that cannot be sustained in the medium to short run, that is to say, inconsistent fixed regimes. This seems to occur since an increase in the U.S. interest rate produces capital outflows from the Latin American region, as found in Calvo, Leiderman, and Reinhart (1993) and Fernández-Arias and Montiel (1996). The de jure regime may try to signal stability as an attempt to control the market instability with mere words. The multinomial logit model corroborates this view: all three estimators are significantly negative, indicating that the likelihood of inconsistent fixed regimes increases when the U.S. interest rate increases.

Openness has the predicted negative sign in the de facto classification (column 2) (that is, economies that are more open prefer a more stable exchange rate). However, the de jure regime (column 1) is significantly positive. According to the multinomial logit results, inconsistent fixed regimes are less likely when *Openness* increases, while consistent fixed and flexible regimes become much more likely. This suggests that open economies are more compatible with macroeconomic strength. The market-based exchange rate tends to float when economies default on their debt (that is, *Default* = 1.573*** in column 2), while the de jure regime keeps pace with the market behavior (that is, *Default* = 0.588*** in column 1). Default definitely decreases the probability of de facto fixed regimes, including both the consistent fixed and inconsistent flexible versions, in relation to inconsistent fixed regimes (that is, *Default* = –0.660*** in column 3; *Default* = –0.493*** in column 4). Consistent flexible regimes become more likely (column 5 is the only positive coefficient), which is congruent with the findings in columns 1 and 2 (that is, de jure and de facto regimes become more flexible and flexible regimes become more consistent).

Now we move on to the issue of endogeneity of regimes around the government change date. The de jure regime does not seem to change in the four quarters leading up to a government change since $GovCh(-3)^q$, $GovCh(-2)^q$, $GovCh(-1)^q$, and $GovCh(0)^q$ are not significant in column 1. This is important since it indicates that de jure regimes are not likely to be strongly affected by the endogeneity of regime announcements.[13] After government changes,

13. In online appendix B, the marginal effects of $GovCh(-3)^q$, $GovCh(-2)^q$, $GovCh(-1)^q$, and $GovCh(0)^q$ are small and insignificant as well.

estimators $GovCh(+1)^q$, $GovCh(+2)^q$, and $GovCh(+3)^q$ in column 1 are significantly positive, indicating that the new government tends to announce more floating.[14] As to the de facto classification, the exchange rate tends to be more flexible both before and after the government change date, since all the government change estimators are significantly positive, with the exception of $GovCh(+4)^q$.

When we use our novel regime classification, the results indicate that no regime is more likely than the baseline (inconsistent de jure fixed regimes) before government changes. This is consistent with the analysis in figure 4 above, in which the regime duration distribution is quite similar for all the regimes in the twelve months before either government changes or a randomly generated month 0. After government changes, the probability of consistent flexible regimes increases relative to inconsistent fixed regimes in the first two quarters: $GovCh(+1)^q = 0.985^{***}$ and $GovCh(+2)^q = 0.838^{**}$. This indicates that the monetary authority announces a flexible regime in the first few months after a government change in an already de facto flexible environment. This is in line with the sharp, sudden drop of inconsistent de jure fixed regimes right after government changes in panel A of figure 4, which contrasts with the behavior after a randomly selected month 0.

The Dynamics of the Real Exchange Rate

Having found no statistical evidence that exchange rate regime announcements vary before government changes, we study the dynamics of the real exchange rate around government changes conditional on consistent/inconsistent de jure regimes. We use a dynamic distributed lag model of the following form:

$$\begin{aligned}(2)\qquad \Delta \ln\left(\mathrm{RER}_{it}\right) = {} & \textstyle\sum_{k=1}^{3} a_k \Delta \ln\left(\mathrm{RER}_{i,t-k}\right) + \mathbf{\Delta W_{it}\boldsymbol{\beta}} + \mathbf{GovCh_{it}\boldsymbol{\delta}} \\ & + \mathbf{GovChFI_{it}\boldsymbol{\delta}_{FI}} + \mathbf{GovChFEI_{it}\boldsymbol{\delta}_{FEI}} \\ & + \mathbf{GovChFEC_{it}\boldsymbol{\delta}_{FEC}} + \varepsilon_{it},\end{aligned}$$

where i and t stand for country and month. The dependent variable is the log difference of the real exchange rate. We control for three distributed lags to

14. These results are in line with Klein and Marion (1997) and Gavin and Perotti (1997) and are also consistent with the pattern found by Blomberg, Frieden, and Stein (2005), where the duration of pegs increases before elections and decreases afterward.

capture persistency.[15] Government change is represented as a matrix of quarterly dummy variables, where $GovCh(\pm l)^q$ takes a value of one if the government change is $\pm l$ quarters away:

$$\mathbf{GovCh_{it}} = \Big[GovCh(-3)^q_{it}\; GovCh(-2)^q_{it}\; GovCh(-1)^q_{it}\; GovCh(0)^q_{it}$$
$$GovCh(+1)^q_{it}\; GovCh(+2)^q_{it}\; GovCh(+3)^q_{it}\; GovCh(+4)^q_{it}\Big].$$

The analogous matrices **GovChFI**, **GovChFEI**, and **GovChFEC** capture the interaction between **GovCh** and inconsistent de jure fixed, inconsistent de jure flexible, and consistent de jure flexible regimes, respectively; the omitted category is consistent de jure fixed regimes. The regime classification used for the entire electoral window is invariant and equal to the classification at the month before the elections, which is typically two or three months before the government change.[16] **W** is a matrix of time-varying controls that attempt to control for the determinants of both exchange rate dynamics and regime announcement. In that regard, we use the same set of variables employed in the estimation of equation 1 to control for the determinants of regime announcement, namely, *Portfolio*, *Foreign.Liab.pc*, *Size, ToT*, *U.S.Interest*, *Openness*, and *Default*. Insofar as an expansion in the size of government will induce an appreciation of the real exchange rate when government demand is biased toward nontradable goods, as stressed by Goldfajn and Valdés (1999), we add government expenditure as a ratio of GDP, *GovSize*. Because we could not corroborate that our regressors produce a cointegrating vector, we estimate the model in first differences, following Cermeño, Grier, and Grier (2010).[17] However, our results do not change significantly when we study equation 2 in levels.[18] Finally, given the possibility of reverse causality, we use one-month lagged values of the variables in **W**. For the variables available at annual frequency that were interpolated using log differences, we use twelve-month lagged

15. Results are totally invariant to the inclusion of one lag. Results with one lag are available on request.

16. Results are virtually unchanged when we use the value six months before elections instead. Results under the latter are available on request.

17. We ran Engle-Granger tests for each country, and in almost all the countries the hypothesis of cointegration was rejected. Only Argentina, Bolivia, Guatemala, Honduras, and Uruguay showed evidence of cointegration at 5 percent significance or higher. Test results are available on request.

18. Results are available on request.

values instead. The rest of the variables are in natural logs except for *Default* and the dummy variables for government change. Results are displayed in table 2. Column 1 shows the results of the estimation of equation 2 for a set of quarterly dummy variables that indicates the proximity of government change date up to four quarters before and after the change. Column 2 replicates column 1 without using covariates, while column 3 replicates column 2 for the extended sample, which includes those countries that we lost owing to covariate limitations.

In column 1 of table 2, the real exchange rate decreases (that is, appreciates) moderately during the last quarter before the government change for inconsistent de jure fixed regimes, but the result is not significant, with $GovChFI(0)^q = -4.832$; after the government change, the real exchange rate depreciates 17 percent in the first quarter, with $GovChFI(+1)^q = 17.312^*$. Linear combination 1, which shows the difference estimator for the two quarters, is not statistically significant, but linear combinations 2, 3, and 4, which progressively extend the window to cover the second, third, and fourth quarters around the government change date, do capture statistically significant depreciation differentials of 16, 11, and 9 percent, respectively; this undoubtedly reflects the significant depreciation during the second quarter, when $GovChFI(+2)^q = 9.469^*$. Column 2, which excludes time-varying covariates, is almost the same as column 1. This indicates that the empirical design produces a plausible exogenous variation of regime adoptions, as shown in figure 4.[19] When we use the extended sample in column 3, the effects drop substantially, but they are all statistically significant.[20]

Real Exchange Rate Misalignments around Government Changes

In the previous section we studied the short- and medium-term dynamics of the real exchange rate (RER) and found that there is a slight and insignificant appreciation quarter to quarter during the year leading up to the government change under inconsistent de jure fixed regimes, followed by a strong and significant depreciation after the change of government. In this section, we explicitly

19. Appendix D1 shows that the results are also robust to allowing for conditional heteroskedasticity.

20. The linear combinations for the countries that are only in the extended sample, though smaller in magnitude, are also positive and statistically significant. For example, linear combination 4 is 9.371** for the reduced sample, 4.932** for the extended sample, and 2.268*** for the extra countries. These regression results are available on request.

TABLE 2. Exchange Rate Variation around Government Changes

Explanatory variable	*Reduced sample, with covariates (1)*		*Reduced sample, no covariates (2)*		*Extended sample, no covariates (3)*	
$GovCh(-3)^q$	−0.358	(0.392)	−0.208	(0.374)	−0.116	(0.286)
$GovCh(-2)^q$	0.417	(0.277)	0.586*	(0.289)	0.362	(0.264)
$GovCh(-1)^q$	−0.540*	(0.281)	−0.328	(0.295)	−0.508**	(0.225)
$GovCh(0)^q$	−0.070	(0.649)	0.222	(0.583)	0.911	(0.790)
$GovCh(+1)^q$	0.697	(0.624)	0.706	(0.642)	0.019	(0.402)
$GovCh(+2)^q$	0.148	(0.281)	0.109	(0.190)	−0.271	(0.283)
$GovCh(+3)^q$	0.028	(0.357)	0.233	(0.352)	0.272	(0.250)
$GovCh(+4)^q$	−0.376	(0.441)	−0.322	(0.419)	0.221	(0.413)
$GovChFI(-3)^q$	−2.647	(2.526)	−2.757	(2.705)	−1.181	(1.512)
$GovChFI(-2)^q$	−2.754	(2.960)	−3.027	(2.592)	−2.015*	(1.113)
$GovChFI(-1)^q$	0.470	(0.833)	−0.130	(0.969)	−0.964	(1.084)
$GovChFI(0)^q$	−4.832	(6.787)	−5.060	(6.503)	−3.416	(2.405)
$GovChFI(+1)^q$	17.312*	(9.345)	16.899*	(8.908)	9.054*	(4.837)
$GovChFI(+2)^q$	9.469*	(5.188)	8.665	(5.227)	2.499	(2.771)
$GovChFI(+3)^q$	0.525	(1.035)	0.389	(1.003)	−0.413	(0.916)
$GovChFI(+4)^q$	0.663	(0.942)	0.556	(0.726)	1.013	(0.996)
$GovChFEI(-3)^q$	0.818	(0.592)	0.897	(0.517)	−0.112	(0.500)
$GovChFEI(-2)^q$	−0.367	(0.447)	−0.360	(0.406)	−0.727***	(0.217)
$GovChFEI(-1)^q$	1.331**	(0.507)	1.061**	(0.373)	0.595*	(0.310)
$GovChFEI(0)^q$	−0.375	(0.669)	−0.613	(0.592)	−1.229	(0.789)
$GovChFEI(+1)^q$	−0.623	(0.694)	−0.554	(0.693)	−0.002	(0.455)
$GovChFEI(+2)^q$	−0.891	(0.957)	−0.300	(0.350)	0.119	(0.324)
$GovChFEI(+3)^q$	−0.636	(0.431)	−0.510	(0.474)	−0.142	(0.357)
$GovChFEI(+4)^q$	0.364	(0.598)	0.529	(0.513)	−0.095	(0.487)
$GovChFEC(-3)^q$	−0.814	(1.068)	−1.020	(1.083)	−0.921	(0.782)
$GovChFEC(-2)^q$	−0.626	(1.328)	−0.813	(1.437)	−0.360	(0.933)
$GovChFEC(-1)^q$	1.545	(1.637)	1.426	(1.579)	1.314	(1.187)
$GovChFEC(0)^q$	0.151	(1.038)	−0.241	(1.077)	−0.203	(1.235)
$GovChFEC(+1)^q$	0.543	(1.572)	0.509	(1.575)	0.897	(1.068)
$GovChFEC(+2)^q$	−0.425	(0.643)	−0.524	(0.606)	−0.522	(0.559)
$GovChFEC(+3)^q$	0.293	(0.652)	−0.112	(0.672)	−0.394	(0.587)
$GovChFEC(+4)^q$	0.843	(0.636)	0.463	(0.684)	−0.520	(0.546)
No. observations	2,236		2,236		4,008	
R^2	0.083		0.056		0.028	
Linear combination 1	22.14	(15.150)	21.96	(14.35)	12.47**	(6.498)
Linear combination 2	15.57**	(7.190)	15.38**	(7.070)	7.967**	(3.650)
Linear combination 3	11.47**	(5.766)	11.39**	(5.605)	5.845**	(2.775)
Linear combination 4	9.433**	(4.761)	9.371**	(4.766)	4.932**	(2.226)

* $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Notes: OLS estimation of equation 2. The dependent variable is $\Delta \ln$ RER. Reduced sample includes Argentina, Bolivia, Brazil, Colombia, Costa Rica, Dominican Republic, Ecuador, Guatemala, Honduras, Mexico, Nicaragua, Panama, Peru, Uruguay, and Venezuela over 1980–99 period. Extended sample also includes Barbados, Chile, El Salvador, Guyana, Jamaica, and Paraguay. *FI*, *FEI*, and *FEC* stand for inconsistent fixed, inconsistent flexible, and consistent flexible regimes; consistent flexible is the omitted category. Nondemocratic episodes are excluded, based on Polity IV Project. Dollarization episodes are also excluded. Controls are used, but not reported (see text). Linear combination $k = 1, 2, 3, 4$: $GovChFI(+1)^q - GovChFI(0)^q$, $\frac{1}{2}(GovChFI(+2)^q + GovChFI(+1)^q - GovChFI(0)^q - GovChFI(-1)^q)$, $\frac{1}{3}(GovChFI(+3)^q + GovChFI(+2)^q + GovChFI(+1)^q - GovChFI(0)^q - GovChFI(-1)^q - GovChFI(-2)^q)$, $\frac{1}{4}(GovChFI(+4)^q + GovChFI(+3)^q + GovChFI(+2)^q + GovChhFI(+1)^q - GovChFI(0)^q - GovChFI(-1)^q - GovChFI(-2)^q - GovChFI(-3)^q)$. Robust standard errors are in parentheses to the right of each estimator.

study the RER misalignment caused by pegging the exchange rate when it is not consistent with the market exchange rate, following the analysis in Goldfajn and Valdés (1999). We control for the stochastic trends of the exchange rate by applying the Hodrick-Prescott filter country by country to each series.[21] The series are then decomposed into two components:

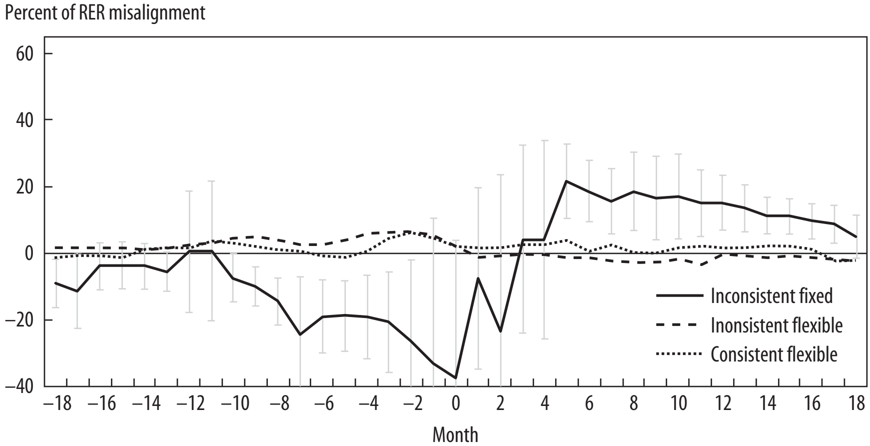
$$\ln\left(\mathrm{RER}_{it}\right) = \ln\left(\mathrm{RER}_{it}\right)_{cycle} + \ln\left(\mathrm{RER}_{it}\right)_{trend}. \tag{3}$$

We identify the trend component as the long-run equilibrium RER and the cycle as departures from that equilibrium. When the cyclical component is negative, the RER is overvalued; when it is positive, it is undervalued. Goldfajn and Valdés (1999) identify four appreciation phases of the RER: history, when the appreciation hits 5 percent; start, when the appreciation hits a threshold (for example, 10 percent, 15 percent); peak, when the appreciation reaches its highest value; and end, when the appreciation is back to the 5 percent history level, which is considered a statistical reversion of the appreciation process. We use this classification to identify when an appreciation represents a significant overvaluation of the exchange rate, in this case, 5 percent and above. The advantage of using logs is that $\ln(\mathrm{RER}_{it})_{cycle}$ already represents the percentage of overvaluation (below the trend) or undervaluation (above the trend). We then estimate the following equation using ordinary least squares (OLS):

$$\begin{aligned}\ln\left(\mathrm{RER}_{it}\right)_{cycle} = {} & \alpha + \sum_{i=0}^{18} GovCh(-i)_{it}^{m} \cdot \beta_{(-i)} + \sum_{i=1}^{18} GovCh(+i)_{it}^{m} \cdot \beta_{(+i)} \\ & + \sum_{i=0}^{18} GovChFI(-i)_{it}^{m} \cdot \beta_{FI(-i)} \\ & + \sum_{i=1}^{18} GovChFI(+i)_{it}^{m} \cdot \beta_{FI(+i)} \\ & + \sum_{i=0}^{18} GovChFEI(-i)_{it}^{m} \cdot \beta_{FEI(-i)} \\ & + \sum_{i=1}^{18} GovChFEI(+i)_{it}^{m} \cdot \beta_{FEI(+i)} \\ & + \sum_{i=0}^{18} GovChFEC(-i)_{it}^{m} \cdot \beta_{FEC(-i)} \\ & + \sum_{i=1}^{18} GovChFEC(+i)_{it}^{m} \cdot \beta_{FEC(+i)} + \varepsilon_{it}.\end{aligned} \tag{4}$$

21. To filter the RER series, we use a smoothing parameter of 129,600, which is the value Ravn and Uhlig (2002) suggest for monthly data. Since the Hodrick-Prescott filter usually introduces a spurious dynamic relation into the series, we also use the Hamilton (2018) filtering technique. The results are qualitatively the same; see online appendix C.

FIGURE 5. Real Exchange Rate Misalignments around Government Changes



Notes: Graphic representation of the estimators of equation 5 for the cyclical component of the RER for the reduced sample, where the RER series is detrended using the Hodrick-Prescott filter (smoothing parameter of 129,600). Results are relative to consistent fixed episodes. Month 0 is the government change month. Vertical bars represent 95 percent confidence intervals of estimators based on robust standard errors for inconsistent fixed estimators around government change date.

In this particular case, we use monthly, rather than quarterly, dummy variables to identify precisely the months at which the overvaluation begins and when it reverts. Afterward, we collapse these monthly dummy variables into six-month periods to observe medium-run misalignments. For the sake of presentation, results of the monthly dummy variables are shown in figure 5, while the half-yearly dummy variables are shown in table 3.

Figure 5 shows that a significant overvaluation occurs only for the inconsistent de jure fixed regime announcement. The 5 percent history threshold is reached at month 10 before the government change, and a peak of 37 percent is reached at the government change month, that is, the history/peak stage lasts ten months, while the peak/end period lasts only three and is mostly completed in the first month. After the government change date, there is a process of undervaluation, which becomes significant at month 5 (undervaluation of 22 percent), but the process reverts smoothly in 14 months, when the RER reaches its equilibrium (that is, back to below 5 percent of undervaluation). Notably, when the exchange rate is overvalued, a quick one-month correction is observed, which indicates that this is achieved through a strong nominal devaluation, as highlighted in Goldfajn and Valdés (1999).

TABLE 3. Real Exchange Rate Misalignments Using Six-Month Dummy Variables

Explanatory variable	*(1)*	
$GovCh(-2)^s$	−1.513	(1.008)
$GovCh(-1)^s$	−2.688***	(0.962)
$GovCh(0)^s$	−3.041***	(0.934)
$GovCh(+1)^s$	0.975	(1.070)
$GovCh(+2)^s$	1.002	(1.035)
$GovCh(+3)^s$	−0.407	(0.926)
$GovChFI(-2)^s$	−4.551**	(2.191)
$GovChFI(-1)^s$	−12.344***	(3.038)
$GovChFI(0)^s$	−25.498***	(5.556)
$GovChFI(+1)^s$	2.981	(6.359)
$GovChFI(+2)^s$	16.292***	(2.226)
$GovChFI(+3)^s$	9.765***	(1.305)
$GovChFEI(-2)^s$	1.733	(1.090)
$GovChFEI(-1)^s$	3.734***	(1.004)
$GovChFEI(0)^s$	5.086***	(1.012)
$GovChFEI(+1)^s$	−0.943	(1.046)
$GovChFEI(+2)^s$	−2.255**	(1.128)
$GovChFEI(+3)^s$	−1.233	(1.021)
$GovChFEC(-2)^s$	0.429	(1.283)
$GovChFEC(-1)^s$	1.664	(1.582)
$GovChFEC(0)^s$	2.831	(2.096)
$GovChFEC(+1)^s$	2.111	(1.710)
$GovChFEC(+2)^s$	1.381	(1.474)
$GovChFEC(+3)^s$	0.449	(1.295)
No. observations	2,127	
R^2	0.119	

* $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Notes: OLS estimation of equation 5 for the cyclical component of RER using six-month dummy variables and the reduced sample, where the RER series is detrended using the Hodrick-Prescott filter (smoothing parameter of 129,600). The dependent variable is $\ln(RER_{it})_{cycle}$. Results are relative to consistent fixed episodes. Nondemocratic episodes are excluded, based on Polity IV Project. Dollarization episodes also excluded. Robust standard errors are in parentheses to the right of each estimator.

However, when the exchange rate is undervalued, as in month +5, the correction takes place smoothly through either a gradual correction of the nominal exchange rate, which corrects the initial overshooting that brought about the undervaluation, or an organized correction of inflation differentials. This difference in the RER reversion in the appreciation and depreciation phases is not treated in Goldfajn and Valdés (1999). Hence, our paper highlights the large asymmetries of the reversions that occur in the overvaluation and undervaluation phases. Furthermore, Goldfajn and Valdés (1999) identify appreciation dynamics without characterizing and describing the context in which these appreciations take place. We identify one particular context where these appreciations occur: elections that coincide with a poor macroeconomic performance.

Finally, table 3 shows the collapsed six-month dummy variables to capture average medium-term behavior. For the inconsistent de jure fixed regimes, the six-month average overvaluation reaches 25 percent in the six months before the government change, which reverts in the first six months after government change, followed by a 16 percent undervaluation in the second six-month period, which then declines to 9 percent in the third period.[22]

Comparison of the 1980–99 and 2000–16 Periods

In this section, we first compare the characteristics of exchange rate regimes for the 1980–99 and 2000–16 periods. We then examine the 2015 Argentine general elections, which share the underlying characteristic of inconsistent de jure fixed regimes, and compare the case study with the econometric findings of the earlier period.

There are eighty-one changes of government in the 1980–99 period and eighty-four in the 2000–16 period. Table 4 shows that de facto flexible regimes are much less common in the recent period, falling from 57 to 30 percent of total cases from one period to the next. While both de jure and de facto classifications are available for the earlier period, only the de facto classification is available for the later period. Nevertheless, we can use the typical characteristics of inconsistent de jure fixed regimes in the earlier period to draw parallels for the more recent period. For 1980–99, two common shared characteristics of inconsistent fixed regimes, where the announced fixed regime does not coincide with the actual policy, are dual/multiple markets and high inflation (more than 10 percent a year). Dual markets and high inflation characterize 81 percent of the inconsistent fixed regime cases in the earlier period, as well as 73 percent of the consistent flexible regime cases, while within the total

22. We also produce both the figure and the table of RER misalignments using the Hamilton (2018) filter. The results, shown in online appendix C, do not change significantly. The results are also robust to allowing for conditional heteroskedasticity; see appendix D2.

TABLE 4. Regime Classification, 1980–2016: Number of Episodes at Government Change Date

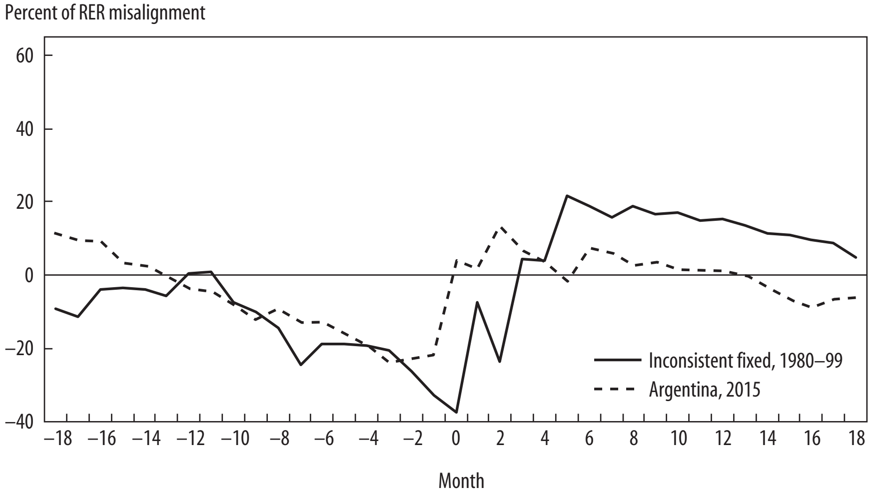
Period and regime	*Total cases*	*Cases with dual markets*	*Cases with annual inflation > 10%*	*Cases with dual markets and annual inflation > 10%*
1980–99 period				
De facto flexible				
Inconsistent de jure fixed	16	16	13	13
Consistent de jure flexible	30	23	27	22
Total de facto flexible	46	39	40	35
De facto fixed				
Inconsistent de jure flexible	24	8	14	6
Consistent de jure fixed	7	2	3	2
Total de facto fixed	31	10	17	8
2000–16 period				
Total de facto flexible	25	1	20	1
Total de facto fixed	49	3	33	2

Note: Dollarization episodes are excluded.

de facto flexible cases, 76 percent involve dual markets and high inflation. This proportion drops to only 4 percent in the later period, and 4 percent represents a single case. Hence, although we cannot observe the IMF de jure classification after 1999, the fact that dual/multiple markets and high inflation become exceptional after 1999 provides indirect evidence that inconsistent fixed regimes are no longer likely.

The only case that we can identify as likely to have been classified as an inconsistent de jure fixed regime—because it shows de facto flexible behavior, dual exchange rates, and high inflation—is the 2015 presidential election of Argentina. The Central Bank of Argentina was applying a crawling peg, devaluing around 1 percent per month from January to November 2015. The incumbent government also announced that the policy would continue the following year, selling huge amounts of future dollar contracts at prices consistent with that crawling peg. Figure 6 compares the real exchange rate misalignment of inconsistent fixed regimes in 1980–99, shown in figure 5, with the real exchange rate misalignment around the government change of Argentina in 2015. As the figure shows, the Argentine currency had an overvaluation of 22 percent in the months leading up to the election, which was corrected suddenly in the election month. The pattern is quite similar to the average trend for inconsistent fixed regimes in the earlier period, where overvaluation peaks at 37 percent in the government change month and is then corrected sharply in about two to three months. We also observe a post-government-change

FIGURE 6. Real Exchange Rate Misalignments: 1980–99 Period versus Argentina 2015



Notes: Graphic representation of the estimators of equation 5 for the cyclical component of the RER for the reduced sample of inconsistent de jure fixed regimes in the 1980–99 period, where the RER series is detrended using the Hodrick-Prescott filter (smoothing parameter of 129,600), and for the Argentina 2015 general elections. Month 0 is the government change month.

undervaluation that is corrected smoothly; for the 1980–99 period, the undervaluation peaks in month +5, while for Argentina 2015, it peaks in month +2.[23]

Conclusion

To explore the behavior of exchange rate policy around elections, we first classified regime announcements using the IMF de jure classification, identifying a regime as inconsistent (broken promises) when it differs from the corresponding Reinhart and Rogoff (2004) de facto classification. We then used ordered logit regressions to study the determinants of both de jure and de facto exchange rate regimes, employing several time-varying controls used in the literature to isolate the impact of dummy variables for government changes (for example, Alesina and Wagner, 2006; Juhn and Mauro, 2002;

23. Online appendix C presents the comparison of the 1980–99 period with Argentina 2015 using the Hamilton (2018) filter rather than the Hodrick-Prescott filter. The results are qualitatively the same.

Levy-Yeyati, Sturzenegger, and Reggio, 2010). We found that de jure regimes do not change in the four quarters leading up to a government change. This is important since it indicates that de jure regimes are not likely to be strongly affected by the endogeneity of regime announcements.[24] Next, we combined the two classifications to study the consistency of the announcement, rather than that of either the announcement or the de facto regime independently from each other. This is something new in the literature on exchange rate regimes.[25] We found that in the first two quarters after government changes, the probability of consistent de jure flexible regimes increases in relation to inconsistent de jure fixed regimes. This indicates that the monetary authority announces flexible regimes the first few months after a government change in an already de facto flexible environment.

We then used this classification to study the dynamics of the real exchange rate around elections conditional on consistent and inconsistent exchange rate regime announcements. We employed a dynamic distributed lag model and a difference-in-differences strategy. This revealed that the pattern found in the earlier political economy literature regarding incumbents who postponed depreciations until the inauguration of the new administration (for example, Cermeño, Grier, and Grier, 2010; Edwards, 1994; Stein and Streb, 2004; Stein, Streb, and Ghezzi, 2005) is specifically due to inconsistent fixed regimes. We found that during inconsistent fixed exchange rate announcements, the devaluation rate is not statistically different from consistent fixed announcements until the government change date, after which it increases and differs from the latter significantly. Although what Alesina and Wagner (2006) call fear of pegging (breaking commitments to pegging and floating more than announced) already shows up in our sample before the end of the incumbent's term, the adjustment of the official exchange rate takes place only *after* the change of government. In other words, part of the broken promise—the devaluation of the official exchange rate—shows up only afterward. One possible interpretation is that sustaining a peg before the government change date can be used as a signal of macroeconomic strength that could increase the probability of being reelected. Exchange rates can be stabilized in the short run by using international reserves and debt. Some incompetent incumbents may

24. In online appendix B, the marginal effects of $GovCh(-3)^q$, $GovCh(-2)^q$, $GovCh(-1)^q$, and $GovCh(0)^q$ are small and insignificant as well.

25. Alesina and Wagner (2006) provide a specific study of inconsistent fixed regimes, which they call fear of pegging. We develop a slightly different classification of the consistency of the announcement and also control for elections.

attempt to mimic competent ones by sustaining the peg announcement before elections (Stein and Streb, 1998, 2004).[26] However, in our sample the postponement of exchange rate adjustments is specifically linked to inconsistent fixed regimes. Hence an additional mechanism is at play: dual markets. Our results on inconsistent fixed regimes thus also suggest the presence of a channel of distributive politics. Specifically, maintaining an "official" appreciated exchange rate before elections hurts the concentrated export sectors to the benefit of the general population that consumes those goods, in particular the median voter. Afterward, the new administration devalues the exchange rate owing to the impossibility (or inconvenience) of sustaining it any longer. This resembles the logic behind the Bonomo and Terra (2005) model, which emphasizes the distributive consequences of appreciated exchange rates, though they do not consider the channel of dual markets. This could be an interesting topic for further research.

Finally, our paper contributes to the literature on real exchange rate appreciations and their reversions. Goldfajn and Valdés (1999) show that real exchange rate appreciations are usually reverted by nominal devaluations rather than through smooth inflation differentials. We identified the episodes of real exchange rate overvaluation as corresponding to inconsistent fixed regime announcements. This starts ten months before the government change date and peaks in the month of government change, with an overvaluation of 37 percent. The overvaluation is mostly reverted in one month through a sudden nominal devaluation. This process leads to a sharp undervaluation of the exchange rate, which is gradually corrected over the course of more than a year. We thus identified a precise timing for the macroeconomic scenario in which exchange rate overvaluation occurs: before the change of government. Additionally, a significant undervaluation takes place in its aftermath, in line with exchange rate overshooting.

26. Following the approach to political budget cycles under asymmetric information in Rogoff and Sibert (1988) and Rogoff (1990), Stein and Streb (1998, 2004) show that a low rate of devaluation can be used before elections by office-motivated incumbents to signal higher competence. In a two-sector model, the postponement of devaluations provokes an appreciated exchange rate (Stein, Streb, and Ghezzi, 2005). In these models, where nominal devaluation acts as a tax on consumption, tax smoothing is optimal from a welfare perspective, but some incumbents are tempted to exploit the trade-off between present and future devaluation for electoral reasons. In a setting with adaptive expectations, van der Ploeg (1989) derives a similar pattern, where the government appreciates the exchange rate before an election, to increase the real income of voters and boost its popularity, and depreciates it afterward. However, his prediction that all incumbents engage in this electoral manipulation is at odds with the evidence.

As to future lines of research, it may be interesting to study how the institutional setup affects the consistency of exchange rate regime announcements. The literature on central bank independence mainly focuses on outcomes like inflation and economic performance (for example, Alesina and Summers, 1993; Garriga and Rodriguez, 2020) or on exchange rate manipulation and volatility (for example, Cermeño, Grier, and Grier, 2010). Higher degrees of central bank independence might increase the likelihood of consistent exchange rate regime announcements (fixed and flexible) during the electoral window and beyond. It may also be interesting to study whether inconsistent fixed regimes lead to a lower probability of reelection and, more generally, whether multiple exchange rate markets affect electoral results.

References

Alesina, Alberto, and Lawrence H. Summers. 1993. "Central Bank Independence and Macroeconomic Performance: Some Comparative Evidence." *Journal of Money, Credit, and Banking* 25 (2): 151–62.

Alesina, Alberto, and Alexander Wagner. 2006. "Choosing (and Reneging on) Exchange Rate Regimes." *Journal of the European Economic Association* 4 (4): 770–99.

Baxter, Marianne, and Alan Stockman. 1989. "Business Cycles and the Exchange Rate Regime: Some International Evidence." *Journal of Monetary Economics* 23 (3): 377–400.

Blomberg, Brock, Jeffry Frieden, and Ernesto Stein. 2005. "Sustaining Fixed Rates: The Political Economy of Currency Pegs in Latin America." *Journal of Applied Economics* 8 (2): 203–25.

Bonomo, Marco, and Cristina Terra. 2005. "Elections and Exchange Rate Policy Cycles." *Economics and Politics* 17 (2): 151–76.

Calvo, Guillermo A., Leonardo Leiderman, and Carmen M. Reinhart. 1993. "Capital Inflows and Real Exchange Rate Appreciation in Latin America: The Role of External Factors." *IMF Staff Papers* 40 (1): 108–51.

Calvo, Guillermo A., and Carmen M. Reinhart. 2002. "Fear of Floating." *Quarterly Journal of Economics* 118 (2): 379–408.

Cermeño, Rodolfo, Robin Grier, and Kevin Grier. 2010. "Elections, Exchange Rates, and Reform in Latin America." *Journal of Development Economics* 92 (2): 166–74.

Edwards, Sebastian. 1994. "The Political Economy of Inflation and Stabilization in Developing Countries." *Economic Development and Cultural Change* 42 (2): 235–66.

Fernández-Arias, Eduardo, and Peter J. Montiel. 1996. "The Surge in Capital Inflows to Developing Countries: An Analytical Overview." *World Bank Economic Review* 10 (1): 51–77.

Garriga, Ana Carolina, and Cesar M. Rodriguez. 2020. "More Effective Than We Thought: Central Bank Independence and Inflation in Developing Countries." *Economic Modelling* 85 (February): 87–105.

Gavin, Michael, and Roberto Perotti. 1997. "Fiscal Policy in Latin America." In *NBER Macroeconomics Annual 1997*, edited by Ben S. Bernanke and Julio Rotemberg, pp. 11–72. Cambridge, MA: MIT Press.

Genberg, Hans, and Alexander K. Swoboda. 2005. "Exchange Rate Regimes: Does What Countries Say Matter?" *IMF Staff Papers* 52 (Special Issue): 129–41.

Goldfajn, Ilan, and Rodrigo Valdés. 1999. "The Aftermath of Appreciations." *Quarterly Journal of Economics* 114 (1): 229–62.

Hamilton, James D. 2018. "Why You Should Never Use the Hodrick-Prescott Filter." *Review of Economics and Statistics* 100 (5): 831–43.

Ilzetzki, Ethan, Carmen M. Reinhart, and Kenneth S. Rogoff. 2019. "Exchange Arrangements Entering the Twenty-First Century: Which Anchor Will Hold?" *Quarterly Journal of Economics* 134 (2): 599–646.

Jeanne, Olivier. 2007. "International Reserves in Emerging Market Countries: Too Much of a Good Thing?" *Brookings Papers on Economic Activity* 1 (2007): 1–79.

Jeanne, Olivier, and Romain Rancière. 2011. "The Optimal Level of International Reserves for Emerging Market Countries: A New Formula and Some Applications." *Economic Journal* 121 (555): 905–30.

Juhn, Grace, and Paolo Mauro. 2002. "Long-Run Determinants of Exchange Rate Regimes: A Simple Sensitivity Analysis." IMF Working Paper 02/104.

Klein, Michael W., and Nancy P. Marion. 1997. "Explaining the Duration of Exchange Rate Pegs." *Journal of Development Economics* 54 (2): 387–404.

Levy-Yeyati, Eduardo, and Federico Sturzenegger. 2005. "Classifying Exchange Rate Regimes: Deeds vs. Words." *European Economic Review* 49 (6): 1603–35.

Levy-Yeyati, Eduardo, Federico Sturzenegger, and Iliana Reggio. 2010. "On the Endogeneity of Exchange Rate Regimes." *European Economic Review* 54:659–77.

Ravn, Morten O., and Harald Uhlig. 2002. "On Adjusting the Hodrick-Prescott Filter for the Frequency of Observations." *Review of Economics and Statistics* 84 (2): 371–76.

Reinhart, Carmen M., and Kenneth S. Rogoff. 2004. "The Modern History of Exchange Rate Arrangements: A Reinterpretation." *Quarterly Journal of Economics* 119 (1): 1–48.

Rogoff, Kenneth S. 1990. "Equilibrium Political Budget Cycles." *American Economic Review* 80 (1): 21–36.

Rogoff, Kenneth S., and Anne Sibert. 1988. "Elections and Macroeconomic Policy Cycles." *Review of Economic Studies* 55 (1): 1–16.

Stein, Ernesto, and Jorge M. Streb. 1998. "Political Stabilization Cycles in High-Inflation Economies." *Journal of Development Economics* 56 (1): 159–80.

———. 2004. "Elections and the Timing of Devaluations." *Journal of International Economics* 63 (1): 119–45.

Stein, Ernesto, Jorge M. Streb, and Piero Ghezzi. 2005. "Real Exchange Rate Cycles around Elections." *Economics and Politics* 17 (3): 297–330.

Van der Ploeg, Frederick. 1989. "The Political Economy of Overvaluation." *Economic Journal* 99 (397): 850–55.

ARTURO ANTÓN
Bank of Mexico, Mexico

ALEJANDRO RASTELETTI
Inter-American Development Bank, United States

Taxing Labor Income in an Economy with High Employment Informality

ABSTRACT This paper develops a static general equilibrium model of occupational choice with heterogeneity in both labor and entrepreneurial skills that generates high levels of employment informality. The model uses a detailed structure of personal income taxes (PITs) and subsidies to formal workers to capture the labor wedges present in many countries. These features enable the model to assess how changes in PITs and subsidies affect labor market outcomes and the government's fiscal accounts. The model is calibrated for Mexico, which, like many developing countries, has high levels of labor informality. The model's simulations shed light on the impact of a series of reforms to PITs and subsidy schemes aimed at increasing labor formality among low-income workers. The results suggest that adjusting the current structure of the formal employment subsidy combined with PIT exemptions for low-income workers could reduce informality while marginally improving the government's fiscal balance.

JEL Codes: H24, H30, J24, J46, O17

Keywords: Informal employment, personal income tax, employment subsidy, fiscal accounts

High levels of informal employment are common in developing countries. Worldwide, almost 70 percent of employment in emerging and developing countries is estimated to be informal, compared with less than 20 percent in advanced economies (ILO, 2018). Numerous studies suggest that high labor taxes may be partially responsible for high levels of informal employment (see, for example, Albrecht, Navarro, and Vroman, 2009; Bosch and Esteban-Pretel, 2012, 2015; Fortin, Marceau, and Savard, 1997; Galiani

ACKNOWLEDGMENTS Antón acknowledges financial support from the Inter-American Development Bank. We thank the editor, Julián Messina; two anonymous referees; and Gabriela Calderón, David Kaplan, Carola Pessino, and Rodrigo Suescún for their valuable comments. We also thank Carlos Puertas and Rodrigo Salinas for outstanding research assistance. The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of Bank of Mexico, the Inter-American Development Bank, its Board of Directors, or the countries they represent.

and Weinschelbaum, 2012; Levy, 2008; Saracoglu, 2008; and Ulyssea, 2018).[1] A study by the Organization for Economic Cooperation and Development, the Inter-American Development Bank, and the Inter-American Center of Tax Administrations (OECD/IDB/CIAT, 2016) estimates that labor taxes in Latin America and the Caribbean are equivalent to 21.7 percent of workers' average incomes. Though below the OECD average of 35.9 percent, the substantial share of labor taxes in average income, especially among low-income workers, may encourage informal employment.

This paper uses a static general equilibrium model with heterogeneous taxes and subsidies to analyze how changes in the labor taxation profile affect informal employment and government finances. The model includes fixed shares of entrepreneurs and workers, who behave rationally. Entrepreneurs and workers are endowed with heterogeneous managerial and labor skills, respectively (compare with Allub and Erosa, 2019; Jovanovic, 1994). Heterogeneous labor skills are treated as the basis for income distribution. The latter property is important, as the distribution of labor taxes and subsidies across formal workers is largely determined by their income level.

To quantitatively assess the role of heterogeneous taxes and subsidies in accounting for informality, the model is calibrated for Mexico, which exhibits the high levels of informal employment typical of many developing countries. According to the International Labour Organization (ILO, 2018), 53.4 percent of total employment in Mexico is informal. However, unlike in many other developing countries, low-income formal workers in Mexico are entitled to a government subsidy that may be credited against their income tax liability. This subsidy is based on income level and is progressive. Understanding how changes in the subsidy affect informal employment and the fiscal accounts in Mexico may offer relevant insights for policymakers in other developing countries.

After calibrating the model, we analyze a series of changes to the personal income tax (PIT) and to the subsidy for formal employment (SUFE). These changes are intended to increase labor formality, especially among low-income workers, without imposing a major fiscal cost. The model suggests that changes to the SUFE and PIT may have large positive effects on labor formalization. Specifically, redesigning the SUFE and including PIT exemptions for low-income workers may boost the rate of labor formality by between 7.0 and

1. Throughout this paper, the term labor taxation is used in a broad sense to refer not only to taxes on personal income but also to mandatory social security contributions.

11.9 percentage points. Moreover, rather than entailing a fiscal cost, these measures improve the government's fiscal balance by 0.34 percent of GDP via their effect on economic formalization. Various sensitivity exercises using alternative values for the model's parameters indicate that these results are robust.

In the model, each entrepreneur receives a managerial ability endowment and runs a single firm. Entrepreneurs use their skills and effective labor to produce a homogeneous good in a competitive context. They must pay a corporate income tax (CIT) and cover the social security contributions (SSCs) of their workers. The CIT is paid in full, and thus these firms are labeled as formal. However, the entrepreneur can hire a wage worker either formally (that is, paying the mandatory SSCs and fringe benefits established by law) or informally (that is, evading such payments). If a worker is hired informally, the entrepreneur faces a probability of being detected and fined by the authorities. This probability is modeled as an increasing function of the firm's size. Therefore, small firms facing a low probability of being fined mostly hire informal workers. For midsize firms, labor is optimally composed of both formal and informal workers. This feature of the model gives rise to an intensive margin of informality, as in Ulyssea (2018). As detailed below, the intensity of labor informality within a firm depends on the level of managerial ability and on the relative costs of formal and informal labor.

The model's workers receive a labor ability endowment and must choose to work on their own or as formal or informal wage employees.[2] Both own-account and informal wage workers pay no taxes on their income and do not contribute to social security, but they receive lump sum transfers from the government. Own-account workers run their own firm without hiring wage workers. Because these firms do not pay taxes, they are classified as informal.[3] By contrast, all formal wage workers pay income taxes and contribute to social

2. Bobba, Flabbi, and Levy (2022) and Narita (2020) also make a distinction between self-employed/own-account and informal wage workers. Self-employment is an important feature of the workforce in developing countries (Gollin, 2008; Perry and others, 2007). In Latin America, it accounts for more than 30 percent of the workforce.

3. Accordingly, formal firms require managerial skills as an input, but informal firms do not. The model structure implies that no entrepreneur operates an informal firm. As a result, there are no informal firms hiring informal wage workers. Evidence from Mexico and Brazil indicates that between 40 and 44 percent of informal employees work for a formal firm, and the remaining work for an informal firm (Samaniego de la Parra, 2017; Ulyssea, 2018). In this regard, the model implies that the intensive margin accounts for 100 percent of informal wage workers, which is at odds with data.

security. If they are low-income earners, they also receive a government subsidy that they can credit against their income tax liability. This scheme of taxes, subsidies, and transfers determines net earnings for workers in each occupation. Given their ability and assessment of the social security services to which they are entitled, workers optimally choose the occupation that yields the highest total earnings.

The large effects on labor formalization found in the quantitative exercises are explained by changes in the net earnings profile of low-income formal workers as a result of variations in the tax and subsidy code. The simulated reform to the SUFE scheme effectively increases the subsidy for formal workers earning between 1.3 and 2.1 times the minimum wage. Meanwhile, the simulated PIT reform provides a tax exemption for formal workers earning up to 1.8 times the minimum wage. In Mexico, approximately 50 percent of employees in the private sector earn up to 2.0 times the minimum wage, and nearly 75 percent of these workers are informal. Therefore, a reform to the tax and subsidy code that increases the earnings of low-income formal workers would significantly alter incentives to formalize.

This paper relates to two broad branches of the literature. The first involves the family of occupational choice models, which has a long tradition in economics (see, for example, Allub and Erosa, 2019; Gollin 2008; Jovanovic, 1994; Lucas, 1978). These models have been used to study how economic agents move between the formal and informal sectors (for example, de Paula and Scheinkman, 2010; Leal, 2014; López, 2017; Pratap and Quintin, 2008; Rauch, 1991). The second branch explores the effects of labor market institutions on informal employment (for example, Albrecht, Navarro, and Vroman, 2009; Bosch and Esteban-Pretel, 2012, 2015; Galiani and Weinschelbaum, 2012; Margolis, Navarro, and Robalino, 2014; Meghir, Narita, and Robin, 2015; Ulyssea, 2010, 2018; Zenou, 2008).

Our model differs from those used in the literature in two important respects. First, we consider heterogeneity in terms of both entrepreneurial and labor abilities. All the works on occupational choice and informality cited above include either heterogeneous entrepreneurial or labor skills, but not both.[4] In our model, the heterogenous distribution of labor skills allows us to build a heterogenous income distribution, which enables us to simulate tax and

4. Jovanovic (1994) and Allub and Erosa (2019) present frameworks with heterogeneity in both managerial and labor skills, and Poschke (2013) uses a model in which both individual ability and firm productivity are heterogenous. However, none of these models incorporates informality.

subsidy policies that depend on workers' income levels. Similarly, heterogeneous entrepreneurial ability plays a key role in determining the intensity of informality within firms in developing countries (Leal, 2014). Our model's second distinguishing feature is its focus on how PIT and subsidy policies for formal workers affect informal employment. The studies cited above examine how labor market policies such as SSCs, unemployment benefits, and restrictions on hiring and firing contribute to informal employment, but none examines how PITs and subsidies for formal workers affect informality.

The rest of the paper is organized as follows. We first describe the model, the data used, and the calibration of the model's parameters. We then use comparative statics to illustrate how changes in tax and subsidy schemes affect labor markets and the public finances. The sensitivity analysis corroborates the robustness of the results. The final section concludes the paper with a brief summary and suggestions for future research.

The Model

The analytical framework is a static general equilibrium model of occupational choice with heterogeneous agents. The economy is composed of two types of agents, entrepreneurs and workers, both of which are independently distributed in fixed proportions. A continuum of managerial and labor abilities is represented by a probability distribution. At the beginning of the period, each entrepreneur is assigned exogenous managerial ability z, which affects the productivity of the firm, while each worker is assigned exogenous labor ability e, which affects her labor earnings. The cumulative distributions of managerial and labor abilities are represented by $\Phi_z(z)$ and $\Phi_e(e)$, respectively.

Each entrepreneur owns a firm that aims to maximize profits based on technology and the structure of taxes and transfers. Firms produce a single good in a competitive context, and each employer hires both formal and informal wage workers in a competitive labor market. When a worker is hired formally, the firm must pay all nonwage labor costs. Alternatively, the firm may avoid these costs by hiring a worker informally. Firms that hire informal workers face a size-dependent probability of being detected and fined by the authorities. All firms pay a corporate income tax (CIT) at a flat rate, which cannot be avoided.

Based on their ability level, workers must select among three possible occupation types: own-account, informal wage employment, or formal wage employment. Workers in the first two occupation types are informal because they pay no taxes or social security contributions (SSCs), and they receive lump sum transfers from the government. Workers in the third occupation type, formal wage employment, must pay personal income taxes (PITs), but they receive social security benefits and, depending on their income level, may receive a government subsidy (namely, the subsidy for formal employment, SUFE). When selecting an occupation, workers compare the amount of labor income they would receive in each of the three occupation types given their skill level, the equilibrium wage, and the structure of taxes and transfers.

Our model distinguishes between formal firms and formal workers: a firm is formal if it pays the CIT, whereas a worker is formal if the employer covers the SSCs. In our model, all entrepreneurs operate formal firms, but own-account workers run informal firms that pay no CIT. They are also informal workers because they do not pay SSCs. All other informal workers are employed by formal firms that do not cover their SSCs. This oversimplification does not allow for informal firms hiring wage workers.

The Entrepreneur's Problem

To produce goods, an employer with ability z must hire wage workers either formally or informally. The relevant input for the firm is effective labor. Let l_F and l_I represent the number of formal and informal workers, respectively. Recalling that e denotes the worker's level of ability, $h_F \equiv el_F$ and $h_I \equiv el_I$ represent the effective labor of formal and informal workers, respectively.

Sorting between employers and wage workers is represented by the function $e = \upsilon(z)$, indicating which employer of ability z is matched with which worker of ability e. We assume positive assortative matching between employers and workers, namely $\upsilon'(z) > 0$. This assumption implies that high-skill employers are matched with high-skill workers, while low-skill employers are matched with low-skill workers.

Technology is represented by the following production function:

$$Y(z) = AzH(z)^{\gamma}, \tag{1}$$

where A is a technology parameter and $\gamma \in (0,1)$ is the Lucas (1978) "span-of-control" parameter. In equation 1, the scale of production (and thus the firm size) increases in relation to managerial ability z. Similarly, $H(z)$ represents the total units of effective labor, as determined according to the following constant elasticity of substitution (CES) function:[5]

(2) $$H(z) = \left\{q(z)h_F^{\psi} + \left[1 - q(z)\right]h_I^{\psi}\right\}^{1/\psi}.$$

In equation 2, the term $q(z)$ determines the relative importance of formal labor in the production process for a given level of managerial ability z. To capture the empirical fact that larger firms in developing countries demand more formal workers (Leal, 2014), we assume that the function $q(z)$ satisfies $q'(z) > 0$. The elasticity of substitution between formal and informal labor in equation 2 is given by $1/(1 - \psi)$, with $\psi < 1$.

Entrepreneurs must pay an output tax at the flat rate τ_Y. They must also cover the wage rate w_F and the corresponding nonwage cost $\tau(z)$ of their formal workers, expressed as a share of the wage cost. Nonwage costs include SSCs, state-level payroll taxes, and fringe benefits. Employers may also receive a tax deduction $D(z)$ per formal worker hired that is proportional to the wage cost.[6] Therefore, $\tau_L(z) \equiv \tau(z) - D(z)$ denotes the cost of hiring a formal worker net of deductions, and the net cost of hiring an effective unit of formal labor may be expressed as $C_F(z) \equiv [1 + \tau_L(z)]w_F$. Alternatively, entrepreneurs may hire workers informally at the wage rate w_I. If the authorities discover that an entrepreneur is hiring workers informally, there is a penalty $\sigma > 1$ on the evaded labor taxes, with no possibility of deduction. Let $V(m) \in [0,1]$ represent the probability of a firm of size m being caught hiring an informal worker with $V'(m) > 0$. This property captures the idea that larger firms face a higher probability of being audited and thus fined by the authorities. Accordingly, the expected cost of hiring an effective unit of informal labor is $C_I(z) \equiv [1 + \sigma V(m)\tau(z)]w_I$.

5. This CES specification is reminiscent of the canonical model of skill differentials developed by Acemoglu and Autor (2011), where a distinction is made between high- and low-skill workers. Equation 2 assumes that effective units of formal and informal labor are imperfect substitutes. Because expression 2 and the assumption $q'(z) > 0$ may be justified by a model where physical capital is more complementary to formal labor than to informal labor, the equation may be interpreted as a reduced-form expression consistent with a capital-skill complementarity model. We thank an anonymous referee for pointing out this interpretation.

6. Typically, nonwage costs and deductions faced by firms are determined as a function of workers' income. In the model, the income span to set taxes and deductions is generated by multiplying the vector of either managerial or labor skills by a scalar. To save on notation, these variables are expressed as a function of ability only. Given the sorting function $e = \upsilon(z)$ between employers and workers, nonwage costs and deductions faced by firms may be expressed in terms of ability z.

Given the above information, the expected net profits for an entrepreneur with ability z may be expressed as:

$$\Pi(z) = (1-\tau_Y)Y(z) - C_F(z)h_F - C_I(z)h_I. \tag{3}$$

Accordingly, employers must choose $\{h_F, h_I\}$ to maximize their expected net profits (equation 3) subject to the technologies represented by equations 1 and 2, taking wages and tax rates as given. After substituting first-order conditions into equation 2, units of effective labor are given by

$$H(z) = \left[(1-\tau_Y)A\gamma z\right]^{\frac{1}{1-\gamma}} C(z)^{\frac{1-\psi}{\psi(1-\gamma)}}, \tag{4}$$

where

$$C(z) = \left[\frac{q(z)}{C_F(z)^{\psi}}\right]^{\frac{1}{1-\psi}} + \left[\frac{1-q(z)}{C_I(z)^{\psi}}\right]^{\frac{1}{1-\psi}}. \tag{5}$$

Adding equation 4 back into the first-order conditions yields the optimal demand for formal and informal labor:

$$h_F(z) = \left[(1-\tau_Y)A\gamma z\right]^{\frac{1}{1-\gamma}} \left[\frac{q(z)}{C_F(z)}\right]^{\frac{1}{1-\psi}} C(z)^{\frac{\gamma-\psi}{\psi(1-\gamma)}}; \tag{6}$$

$$h_I(z) = \left[(1-\tau_Y)A\gamma z\right]^{\frac{1}{1-\gamma}} \left[\frac{1-q(z)}{C_I(z)}\right]^{\frac{1}{1-\psi}} C(z)^{\frac{\gamma-\psi}{\psi(1-\gamma)}}. \tag{7}$$

Equations 6 and 7 yield the optimal h_F/h_I ratio. Because the effective labor of formal and informal workers may be rewritten as $h_F \equiv el_F = \upsilon(z)l_F$ and $h_I \equiv el_I = \upsilon(z)l_I$, the ratio of formal to informal workers is given by

$$\frac{l_F}{l_I} = \left\{\left[\frac{q(z)}{1-q(z)}\right]\frac{\left[1+\sigma V(m)\tau(z)\right]w_I}{\left[1+\tau_L(z)\right]w_F}\right\}^{\frac{1}{1-\psi}},$$

which indicates that the ratio increases with firm size m, given the assumption $V'(m) > 0$. This pattern is consistent with data for developing countries, where smaller firms are more likely to hire informal workers relative to larger firms, but larger firms still employ a substantial share of all informal workers (Leal, 2014).

The Worker's Problem

As described above, each worker is assigned exogenous labor ability e. When workers choose to become formal wage employees, they receive wage earnings represented by the function $W_F(w_F, e)$ and may be entitled to a subsidy $S(e)$ from the government. They must also pay PIT in the amount of $\tau_W(e)$. Accordingly, their after-tax income $I_F(e)$ is

$$I_F(e) = W_F(w_F, e) + S(e) - \tau_W(e). \tag{8}$$

Formal workers are automatically enrolled in the social security system. If $\tau_{SS}(e)$ denotes the tax rate on SSCs, the contribution paid by an employer for a worker with ability e is $\tau_{SS}(e)W_F(w_F, e)$. Formal workers are entitled to receive social security services such as health care and pensions, but not all services may be fully valued by workers (see Summers, 1989). Let $\beta_F > 0$ denote the valuation made by formal workers of such services. Therefore, the monetized value of social services is expressed as $\beta_F\tau_{SS}(e)W_F(w_F, e)$.[7] Formal workers also receive fringe benefits, which are denoted as a fraction κ of wage earnings $W_F(w_F, e)$. The workers' valuation of these benefits is expressed by parameter $\beta_S > 0$.

Based on the above specifications, the net earnings of formal wage employment for a worker with ability e may be expressed as follows:

$$E_{W,F}(e) = I_F(e) + \left[\beta_F\tau_{SS}(e) + \beta_S\kappa\right]W_F(w_F, e), \tag{9}$$

where $I_F(e)$ is given by equation 8. Each worker must compare these earnings to those generated by other occupation types.

INFORMAL WAGE WORKERS. The wage earnings of an informal wage worker are represented by the function $W_I(w_I, e)$. As noted above, informal workers pay no PIT, receive no subsidy $S(e)$, and are not entitled to social security or

7. When $\beta_F < 1$, SSCs are a tax in net terms. See Summers (1989).

other nonwage employment benefits. However, they receive a noncontributory social security transfer T_{NC} from the government. The valuation of such transfers by workers is captured by parameter $\beta_I > 0$.

Therefore, the total earnings of an informal worker $E_{W,I}(e)$ are given by

$$E_{W,I}(e) = W_I(w_I, e) + \beta_I T_{NC}. \tag{10}$$

Empirical evidence suggests that the returns to education are higher for formal workers than for informal workers (see, for example, Gong and van Soest, 2002; Günther and Launov, 2012). If education levels efficiently signal worker ability, the earnings function of formal workers should exhibit higher returns to scale in ability e relative to the earnings function of informal workers. For simplicity, the earnings function $W_F(w_F, e)$ in equation 8 is set to exhibit constant returns to scale: $W_F(w_F, e) = w_F e$. Accordingly, the earnings function of informal workers in equation 10 is determined by $W_I(w_F, e) = w_I e^{\alpha}$ with parameter $\alpha \in (0,1)$.

OWN-ACCOUNT WORKERS. In our model, own-account workers produce the same goods as entrepreneurs but use slightly different technology, which is represented by the production function $Y_O = A_O h_O^{\gamma_O}$. In this equation, A_O is a technology parameter, $\gamma_O \in (0,1)$ captures the returns to scale in production, and h_O denotes effective units of labor given by $h_O \equiv e l_O$. Because own-account workers pay no taxes and make no SSCs, profits Π_O (before transfers) may be simply written as $\Pi_O(h_O) = A_O h_O^{\gamma_O}$.

Own-account workers are also recipients of noncontributory social security transfers T_{NC}. Assuming that $\beta_I > 0$ captures the valuation of such transfers, own-account earnings are written as

$$E_O(e) = A_O h_O^{\gamma_O} + \beta_I T_{NC}. \tag{11}$$

THE WORKER'S OCCUPATIONAL CHOICE PROBLEM. Having explained the earnings of each type of worker, we now define the occupational choice problem of a worker with ability e. In general, this problem is written as

$$\max_c u(c)$$

subject to:

$$c = I(e),$$

where c denotes consumption and $I(e)$ is the income of a worker with ability e, defined by the following equation:

$$I(e) = \max\left\{E_O(e), E_{W,I}(e), E_{W,F}(e)\right\}. \tag{12}$$

The terms shown on the right-hand side of equation 12 are specified by equations 9, 10, and 11. To simplify, the utility function $u(c)$ is assumed to be linear in consumption, and thus $u(c) = c$. As a result, the utility of each occupation is equivalent to the earnings received (compare with D'Erasmo, Moscoso, and Senkal, 2014; Galiani and Weinschelbaum, 2012).

Based on this framework, we define the sets of own-account ($\mathcal{A}$), informal wage ($\mathcal{B}$), and formal wage workers ($\mathcal{C}$) as follows:

$$\mathcal{A} = \left\{e \middle| I(e) = E_O(e)\right\}; \tag{13}$$

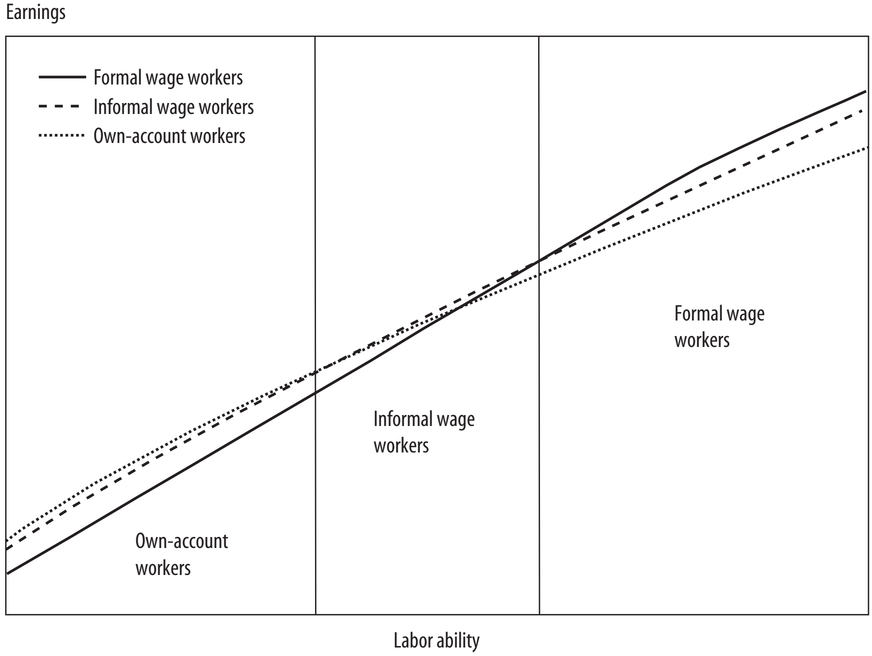
$$\mathcal{B} = \left\{e \middle| I(e) = E_{W,I}(e)\right\}; \tag{14}$$

$$\mathcal{C} = \left\{e \middle| I(e) = E_{W,F}(e)\right\}. \tag{15}$$

For illustrative purposes, figure 1 shows a hypothetical earnings profile for each type of worker and the corresponding occupational choice made as a function of labor ability e. In this case, own-account employment provides the highest earnings for less-skilled workers. At moderate levels of labor skills, the worker optimally chooses informal wage employment, and when labor skills are sufficiently high, the worker chooses formal wage employment. Equations 9, 10, and 11 show that changes in the structures of taxes, subsidies, and transfers may affect occupational choices.

The reform simulations presented later in the paper involve changes in the subsidy and income tax profiles $S(e)$ and $\tau_W(e)$ in equation 8, based on the distribution of managerial and labor abilities. Such changes affect after-tax income and thus the net earnings profile $E_{W,F}(e)$ of formal workers. Consequently, occupational choices may also change. For example, an increase in the subsidy $S(e)$ to low-income formal workers would cause a corresponding increase in net earnings $E_{W,F}(e)$, incentivizing those workers who were initially indifferent between formal and informal wage employment to prefer the former. As a result, the share of formal workers in the economy would

FIGURE 1. Earnings and Occupational Choices



Source: Authors' elaboration.

increase. Reducing income taxes on low-income workers would yield a similar outcome. While general equilibrium effects must also be incorporated into the analysis, changes in $S(e)$ or $\tau_W(e)$ or both will drive the results in the simulations below.

Now let $L_O = \int_A d\Phi_e(e)$, $L_I = \int_B d\Phi_e(e)$ and $L_F = \int_C d\Phi_e(e)$ denote the total number of own-account, informal wage workers, and formal wage workers, respectively. Given that the total number of workers $\bar{L}$ is fixed, the following must hold:

$$L_O + L_I + L_F = \bar{L}. \tag{16}$$

Equation 16 shows that changes to the fiscal structure do not alter the number of workers $\bar{L}$. However, such changes may affect the relative share of workers in each type of occupation.

Equilibrium

In equilibrium, the demand for informal wage workers (measured in units of effective labor) must equal their supply. The same is true for formal wage workers. The labor supply of these two occupation types is determined by the occupational choice problem described above. Therefore, equilibrium conditions for formal and informal wage workers may be expressed as follows:

$$\int_z h_I\left(z, w_F^*, w_I^*\right) d\Phi_z\left(z\right) = \int_{\mathcal{B}} ed\Phi_e\left(e\right); \tag{17}$$

$$\int_z h_F\left(z, w_F^*, w_I^*\right) d\Phi_z\left(z\right) = \int_C ed\Phi_e\left(e\right). \tag{18}$$

Accordingly, equations 17 and 18 solve for equilibrium wages $\{w_F^*, w_I^*\}$.

Calibration

In this section, we calibrate the model using data for Mexico to quantitatively assess how changes in PIT and subsidies to formal workers may affect formal employment and the government's budget balance. According to the ILO (2018), informal employment accounts for 53.4 percent of total employment in Mexico, broadly in line with the average for Latin American and Caribbean economies.

Our quantitative exercise incorporates detailed information on Mexico's PIT and SUFE schemes. The Mexican PIT is progressive, with statutory marginal tax rates starting at 1.92 percent and gradually increasing to a maximum rate of 35 percent. As noted in the introduction, the SUFE is a progressive subsidy provided to formal low-income workers to decrease their income tax burden. Further information on the PIT and SUFE schemes can be found in online appendix A.[8]

The data sources used to calibrate the model are detailed in online appendix B.[9] In 2017, CIT and PIT revenues each amounted to 3.1 percent of GDP, and transfers via the SUFE were equivalent to 0.2 percent of GDP. Between

8. Supplementary material for this paper is available online at http://economia.lacea.org/contents.htm. The model is also calibrated to replicate the Mexican SSC scheme, in which contributions are partially income-based and benefits include health care, pensions, life insurance, housing, and day care. For a thorough description of Mexico's SSC system, see Levy (2008).

9. See IMSS (2018), INEGI (2013, 2014, 2017), and Ministry of Finance (2017a, 2017b, 2017c).

2003 and 2016, SSCs from workers and employers averaged 3.1 percent of GDP, but no data are available on their relative shares. The Mexican government also finances social security for formal workers, and in 2017 its contribution was valued at 0.53 percent of GDP. By law, only a fraction of SSCs must be allocated to health insurance, and in 2017 these contributions fell short of the government's total health expenditures. Therefore, potential increases in formal employment imply additional financial commitments by the government, which we estimate at 13,503 Mexican pesos (MXN) annually per formal worker, based on the official data. These commitments are referred to as extra operating expenditures in the analysis below.

Labor market information is provided by Mexico's National Occupation and Employment Survey (*Encuesta Nacional de Ocupación y Empleo*, ENOE). All calculations exclude public sector employment (encompassing employment by government agencies, state-owned enterprises, and public institutions) because government workers have their own social security scheme and receive benefits that are not comparable to those of private sector workers. The ENOE distinguishes between workers who are affiliated with the SSC scheme (that is, formal workers) and those who are not (that is, informal workers). Employers account for 4.8 percent of total employment, while own-account (26.3 percent), informal wage (39.5 percent), and formal wage (29.4 percent) workers make up the remaining 95.2 percent. Formal wage workers receive an average net wage of MXN 7,447 per month, and informal wage workers receive an average net wage of MXN 4,344 per month at 2018 prices. Though not required for the calibration process, the average earnings of entrepreneurs and own-account workers are also reported.

Functional Forms

The model requires specifying the functional forms for the distribution of skills, the weight of formal workers in the production function, and the probability of detection by the authorities. For the first case, labor ability e is described in terms of a truncated log-normal distribution with mean μ_E, variance σ_E^2, and support $[\underline{e}, \overline{e}]$. Entrepreneurial ability z is defined by a truncated log-normal distribution with support $[\underline{z}, \overline{z}]$, mean μ_Z, and variance σ_Z^2.

For the function $q(z)$ included in equation 2, the following specification is adopted:

$$q(z) = 1 - \exp\left[-\left(\frac{z - \underline{z}}{\lambda}\right)^{\zeta}\right], \tag{19}$$

where $\lambda > 0$ is a scale parameter, while $\zeta > 0$ is a shape parameter. This expression is a variant of the Weibull cumulative distribution function and is sufficiently flexible depending on the values of λ and ζ. Assuming that firm size is proportional to z, the probability of detection $V(m)$ may be expressed as $V(m(z))$. For simplicity, the function $V(m(z))$ is set to depend linearly on $q(z)$.

Parameter Values

The model uses three groups of parameters. The first group reflects the current structure of the PIT, SUFE, and SSC schemes, which are defined by 110 parameters set according to their 2018 values. The second group includes twelve parameters related to technology, preferences, transfers, and the distribution of labor ability. Several parameters within this group are selected to determine the earnings profile of own-account, informal wage, and formal wage workers. Others are fixed according to the available data or values used in the literature. Without further evidence from either the data or the literature, the remaining parameters are set a priori and are subjected to a sensitivity analysis described below.

For technology parameters belonging to the second group, returns to scale in the production function of the entrepreneur, γ, are set at 0.76 following Leal (2014). For simplicity, parameter γ_O is also set at 0.76. The technology level of own-account workers, A_0, is set at 5,574, which yields a reasonable monthly earnings estimate for this occupation type.[10] A value of 0.86 is assigned to parameter α, which measures returns to scale for the skill levels of informal wage workers. Fixing the values for these three technology parameters helps determine the earnings profile of the own-account and informal wage workers not related to lump sum transfers (see equations 10 and 11). The parameter linked to the elasticity of substitution between formal and informal effective labor (ψ) is set at 0.9, which implies a relatively high value for the elasticity of substitution. In the absence of further evidence on α and ψ, alternative values are considered in the sensitivity analysis. When the authorities discover that a firm has evaded SSCs, it is assumed that the firm must cover the evaded

10. In the data, own-account workers earn slightly more, on average, than informal wage earners (see table 4). The two revenues are difficult to replicate in the model simultaneously as the people with the lowest labor skills in the model (and thus with the lowest average earnings) are own-account workers, while those with moderate labor skills are informal wage workers (see figure 1). Therefore, the model is calibrated such that the average income of informal wage earners replicates the data (see table 2). In contrast, the value of parameter A_0 yields an average income of MXN 2,766 per month for own-account workers, which is a reasonable amount, though it falls significantly below the figure observed in the data (see table 4).

TABLE 1. Parameter Values for the Second Group

Parameter	*Value*	*Parameter*	*Value*
γ	0.76	β_F	0.30
γ_0	0.76	β_S	0.90
A_0	5,574	β_I	0.85
α	0.857	T_{NC}	948
ψ	0.90	$\underline{e}$	0.14
σ	1.50	$\overline{e}$	12.00

Source: Authors' elaboration.

amount plus a fine equivalent to 50 percent of that amount.[11] Therefore, we set $\sigma = 1.5$.

Some preference parameters reflect workers' valuation of social security and fringe benefits. For formal employees, we assume that $\beta_F = 0.30$, meaning that workers value only 30 percent of the benefits associated with mandatory SSCs. Similarly, the valuation of fringe benefits, β_S, is fixed at 0.90. The valuation of lump sum transfers to informal workers, β_I, is set at 0.85. As the values for these parameters may be controversial, the sensitivity analysis uses alternative specifications.

The value of government transfers to informal workers, T_{NC}, is estimated at MXN 948 per month, based on Antón and Hernández (2017), and adjusted for inflation using 2018 prices. Compensation for labor ability is defined by $\underline{e} = 0.14$ and $\overline{e} = 12$. These values allow for gross labor incomes ranging from just over MXN 850 to MXN 73,500 per month. Parameter values of the second group are reported in table 1.[12]

The third group includes the remaining ten parameters, which are simultaneously calibrated and for which there are no direct references in the literature. These parameters are related to technology (A, λ, and ζ), fiscal policy (τ_Y), and the distribution of abilities (μ_E, σ_E^2, μ_Z, σ_Z^2, $\underline{z}$, and $\overline{z}$). Since we have ten unknowns, we set ten relevant moments from the theoretical model to match

11. Mexico's social security law establishes fines of between 40 and 100 percent of the amount evaded, based on the severity of the offense. Firms must also cover the evaded amount plus the forgone interest. See Levy (2008) for a discussion.

12. The model is also calibrated to replicate the ratio of fringe benefits over production reflected in the data. In addition, payroll taxes at the state level are set to 2 percent of the wage rate (compare with Antón, Hernández, and Levy, 2012). Recall that these two elements are part of a firm's formal nonwage costs $\tau(e)$. Finally, the model estimates are adjusted to replicate the government's fiscal accounts as a share of GDP and the employers' average earnings under the benchmark scenario.

TABLE 2. Moments Reflected in the Data and Model

Moment	*Data*	*Model*
Share of own-account employment	0.263	0.263
Share of formal wage employment	0.294	0.294
Average income of formal wage workers (pesos per month)	7,447	7,469
Average income of informal wage workers (pesos per month)	4,344	4,419
Share of workers earning up to 1 minimum wage	0.124	0.124
Share of workers earning 1 to 2 times the minimum wage	0.385	0.385
Share of workers earning 5 to 10 times the minimum wage	0.048	0.048
Share of workers earning more than 10 times the minimum wage	0.012	0.012
Share of formal workers earning up to 1 minimum wage	0.004	0.004
Corporate income tax collection (% of GDP)	3.100	3.110

Source: Authors' elaboration.

the data. Moments used for the calibration exercise are reported in table 2. The moments chosen are associated with the relative shares of occupation types, the average income of wage workers, the earnings distribution, and the tax revenue generated by CIT. Given that the earnings profiles of own-account and wage workers are previously determined, parameters within this group can be calibrated to match the share of each occupation type, the average income of wage workers, and the earnings distribution.

The calibration is performed simultaneously because a change in the value of a given parameter affects two or more moments in the model. Nevertheless, some parameters are more useful than others to match specific moments in the data. For example, the distribution parameters μ_E and μ_Z are particularly useful for matching the average income of formal and informal wage workers. Similarly, parameters, σ_E^2 and σ_z^2 are useful for replicating the shares of own-account and formal wage workers. Boundary parameters $\underline{z}$ and $\overline{z}$ are appropriate for matching the share of workers earning up to the minimum wage and more than ten times the minimum wage, respectively.[13] On the other hand, the share of workers earning one to two and five to ten times the minimum wage and the share of formal workers earning up to the minimum wage are matched with the scale and shape parameters, λ and ζ, and the technology parameter, A. Finally, the tax parameter τ_Y is calibrated to replicate CIT revenue as a share of GDP.

13. The difficulties of the log-normal distribution for replicating both the upper and lower tails of actual income distributions are well known (Dagum, 1977). Numerical simulations show that increasing the number of parameters in a generalized version of the log-normal distribution improves the fitness to the data (McDonald and Ransom, 2008). In our case, the truncation of the log-normal distribution allows for a better calibration of the model to the data in both tails.

Once all parameter values are set, equilibrium wages w_F^* and w_I^* solve for equilibrium conditions 17 and 18. Table 2 illustrates how well the model replicates the data.[14] Parameter values obtained under the benchmark calibration are as follows: $A = 27{,}448$, $\lambda = 7.14$, $\zeta = 0.15$, $\tau_Y = 0.08$, $\mu_E = -0.34$, $\sigma_E^2 = 0.25$, $\mu_Z = -1.12$, $\sigma_Z^2 = 0.26$, $\underline{z} = 0.21$, and $\overline{z} = 11.17$.

Reform Simulations

Having calibrated the model to replicate key aspects of the Mexican economy, we now analyze how changes to the SUFE and PIT schemes would affect labor informality and the fiscal accounts.[15] This section presents a series of comparative static exercises designed to elucidate the relevant policies. It is especially critical to understand that the SUFE and PIT are not equivalent. As explained in online appendix A, the SUFE is a transfer to low-income formal workers based on their gross income. The SUFE does not affect the tax base for the workers' PIT, and workers can credit the SUFE against their tax liability. Consequently, the SUFE and PIT may have quantitatively different effects on workers' occupational decisions.

Simulations of an Alternative SUFE Policy

The SUFE in Mexico is granted as a function of gross income to reduce low-income workers' personal income taxes. The scheme is progressive in that the subsidy increases as the worker's income decreases. For illustrative purposes, table 3 presents the SUFE scheme in place in 2018, with the lower and upper bounds of gross monthly income defined by law. For example, if a worker earns MXN 6,500 per month, the SUFE granted amounts to

14. From a technical view, the numerical solution to the nonlinear system above is generally non-unique. We therefore tried alternative initial parameter values to find the best fit to the data. However, the calibration of the income distribution is far from perfect. The fraction of workers earning between two and three times the minimum wage is underestimated by two percentage points, implying that the earnings distribution between three and five times the minimum wage is overestimated by the same amount.

15. The simulation exercises generate a change in the government's budget balance in all cases. To generate a policy that is balance neutral, the government could implement lump sum transfers (taxes) to all workers in the event of an increase (decrease) in the budget balance. Under such a policy, the set of occupational choice allocations described by equations 12–15 would not change because all workers would receive the same lump sum transfer (alternatively, pay the same lump sum tax). Of course, such a policy would change both workers' earnings and utilities.

TABLE 3. SUFE Scheme, 2018
Monthly income in pesos

Lower bound	Upper bound	Subsidy
0.01	1,768.96	407.02
1,768.97	2,653.38	406.83
2,653.39	3,472.84	406.62
3,472.85	3,537.87	392.77
3,537.88	4,446.15	382.46
4,446.16	4,717.18	354.23
4,717.19	5,335.42	324.87
5,335.43	6,224.67	294.63
6,224.68	7,113.90	253.54
7,113.91	7,382.33	217.61
7,382.34	Onward	0

Source: Ministry of Finance.

MXN 253.54. Workers who earn more than MXN 7,382.33 per month have no right to the SUFE. Further explanation of this scheme can be found in online appendix A.

This section analyzes three potential changes to the SUFE: (1) eliminating the policy; (2) switching to a uniform transfer of MXN 400 per month to all formal wage workers regardless of income level; and (3) altering the benefit amount and the eligibility threshold. As shown below in detail, our findings suggest that eliminating the SUFE would reduce the formality rate by six percentage points, with an adverse overall impact on the fiscal accounts owing to rising informality. By contrast, switching to a uniform MXN 400 transfer to all formal workers would increase the formality rate by nearly three percentage points, but this improvement would come at a significant fiscal cost. Finally, altering the benefit amount and the eligibility threshold would increase the formality rate by 2.4 percentage points while yielding a modest improvement in the fiscal accounts.

Table 4 illustrates the effects of changes to the SUFE scheme on labor market outcomes, net incomes, the fiscal accounts, and the burden of PIT and SSCs. The table shows the baseline calibration of the different variables of interest, which can be compared to the values in the data. The definition of the budget balance used in the section on fiscal accounts corresponds to the public revenues and expenditures included in the model, not to the government's actual fiscal balance, which encompasses all public revenues and expenditures. The section on the tax burden reports on how PIT and SSC revenues are distributed between wage workers and employers.

TABLE 4. Subsidy for Formal Employment (SUFE): Reform Simulations

			Reform model		
Variable	*Data*	*Baseline model calibration*	*Elimination of subsidy to formal workers*	*Uniform subsidy of MXN 400 to all formal workers*	*Limited subsidy to formal workers*
Occupation (as a share of employment)					
Total informal	0.658	0.658	0.720	0.630	0.634
Own-account	0.263	0.263	0.312	0.236	0.267
Informal wage	0.395	0.395	0.408	0.394	0.367
Formal wage	0.294	0.294	0.232	0.322	0.318
Employers	0.048	0.048	0.048	0.048	0.048
Average net income (pesos per month)					
Total wage	5,668	5,720	5,892	5,720	5,737
Formal	7,447	7,469	8,049	7,494	6,918
Informal	4,344	4,419	4,664	4,268	4,715
Own-account	4,762	2,766	2,887	2,695	2,778
Employers	12,817	12,817	12,407	12,995	12,805
Fiscal accounts (% of GDP)					
Wage workers					
(A) Income tax	3.10	3.10	3.07	2.11	3.49
(B) SSC	n.a.	0.29	0.24	0.31	0.32
Employers					
(C) Income tax	3.10	3.10	3.33	3.00	3.11
(D) SSC	n.a.	2.81	2.34	3.01	3.04
Government: Contributory SS					
(E) SS revenue (B + D)	3.10	3.10	2.59	3.32	3.36
(F) SS expenditures	0.53	0.53	0.44	0.57	0.57
(G) Extra operating expenditures	. . .	. . .	−0.11	0.08	0.07
(H) Balance (E − F − G)	2.57	2.57	2.26	2.67	2.71
Government: Other					
(I) Income tax revenues (A + C)	6.20	6.20	6.40	5.10	6.60
(J) Subsidy to formal employment	0.20	0.20	0.00	0.59	0.19
(K) Noncontributory SS	1.70	1.70	1.92	1.61	1.64
(L) Balance (I − J − K)	4.30	4.30	4.48	2.91	4.78
Government: Total					
(M) Revenue (E +I)	9.30	9.30	8.99	8.42	9.96
(N) Expenditures (F + G + J + K)	2.43	2.43	2.25	2.85	2.47
(O) Budget balance (M − N)	6.87	6.87	6.74	5.58	7.49
Tax burden (%)					
Income tax					
Wage workers	50.0	50.0	48.0	41.2	52.8
Employers	50.0	50.0	52.0	58.8	47.2
SSC					
Wage workers	n.a.	9.4	9.5	9.4	9.4
Employers	n.a.	90.6	90.5	90.6	90.6

Source: Authors' elaboration.
. . . Not applicable.
n.a. Not available.

In the third column of table 4, the SUFE is eliminated for all formal wage workers, which causes the formality rate to drop from 29.4 to 23.2 percent, as some lower-income workers who currently benefit from the SUFE see their incentives shift in favor of informality. The exit of the lowest-skilled, lowest-paid workers from the formal sector increases the average wage for formal workers, w_F^*, to MXN 8,049 per month. Meanwhile, those previously formal workers become the most-skilled, highest-paid members of the informal workforce, which raises the average income of informal wage and own-account workers. However, the average earnings of employers decrease as average wages rise, suggesting that they benefit indirectly from the SUFE.

The decline in formality boosts CIT revenue because wages paid to formal workers are tax deductible for the employer, whereas the wages of informal workers are not.[16] Nevertheless, total fiscal revenue (row M) falls owing to the drop in both PIT (row A) and worker/employer SSCs (row E). Government spending decreases as the SUFE disappears (row J) and rising informality reduces contributory social security expenditures (row F), but the increase in informality also increases noncontributory social security expenditures (row K). In net terms, the budget balance deteriorates, falling from 6.87 to 6.74 percent of GDP. Overall, the elimination of the SUFE reduces the formality rate by six percentage points while marginally worsening the fiscal balance.

The fourth column in table 4 shows the effects of transforming the SUFE into a uniform transfer of MXN 400 per month to all formal workers. This change increases the formality rate by 2.8 percentage points to 32.2 percent of total employment. The uniform SUFE decreases the PIT liability of all formal workers, increasing their average net income while simultaneously reducing PIT revenue from 3.1 to 2.1 percent of GDP. As PIT revenue falls, the share of the CIT in total revenue rises from 50 to 58.8 percent, raising the tax burden on employers relative to workers. The increase in formalization has a positive fiscal impact, but the cost in forgone PIT revenue outweighs this effect. After the other fiscal implications have been accounted for, the budget balance deteriorates from 6.87 to 5.58 percent of GDP.

The last column in table 4 simulates changes to the SUFE designed to enhance its positive impact on both the formality rate and the fiscal balance. Under these changes, a uniform employment subsidy of MXN 400 per month

16. The elimination of the SUFE entails two conflicting effects on the total cost of hiring formal workers. On the one hand, lower earnings for formal workers cause the formal labor supply to fall and the equilibrium wage to increase. On the other hand, this causes the demand for formal work to decrease in equilibrium. In this specific case, the total cost decreases on average, which indicates that the fall in formal work is more significant than the wage increase.

TABLE 5. PIT Schemes
Monthly income in pesos

Lower bound	*Upper bound*	*Fixed amount*	*Rate (%)*
A. 2018 PIT table			
0.01	578.52	0.00	1.92
578.53	4,910.18	11.11	6.40
4,910.19	8,629.20	288.33	10.88
8,629.21	10,031.07	692.96	16.00
10,031.08	12,009.94	917.26	17.92
12,009.95	24,222.31	1,271.87	21.36
24,222.32	38,177.69	3,880.44	23.52
38,177.70	72,887.50	7,162.74	30.00
72,887.51	97,183.33	17,575.69	32.00
97,183.34	291,550.00	25,350.35	34.00
291,550.01	Onward	91,435.02	35.00
B. PIT table under reforms			
0.01	4,910.18	0.00	0.00
4,910.19	8,629.20	0.00	10.88
8,629.21	10,031.07	404.63	16.00
10,031.08	12,009.94	628.93	17.92
12,009.95	24,222.31	983.54	21.36
24,222.32	38,177.69	3,592.10	23.52
38,177.70	72,887.50	6,874.40	30.00
72,887.51	97,183.33	17,287.34	32.00
97,183.34	291,550.00	25,062.00	34.00
291,550.01	Onward	91,146.67	35.00

Sources: Ministry of Finance and authors' elaboration.

is provided to all formal wage workers earning up to MXN 4,910 per month. The eligibility ceiling for the maximum subsidy corresponds to the current upper bound of the second income bracket of the PIT scheme (see table 5).[17] Among formal wage workers earning more than MXN 4,910 per month, the SUFE decreases linearly until it reaches zero for workers with incomes of MXN 7,410 per month.[18] Overall, these changes shift the distribution of SUFE benefits toward lower-income formal workers.

These changes increase the formality rate while improving the fiscal balances. The positive effect on formality is similar to that observed under the uniform SUFE transfer. However, the fiscal balance changes markedly.

17. This amount represents 1.8 times the minimum wage for 2018. According to the ENOE, approximately 50 percent of Mexican employees in the private sector earn up to twice the minimum wage.

18. The gradual reduction of the SUFE is designed to ameliorate disincentives to formality generated by an abrupt elimination of the subsidy.

On one hand, higher levels of formality boost the collection of PIT and SSCs while leaving CIT revenues broadly unchanged. On the other hand, the fiscal cost of the SUFE declines relative to the baseline, as does spending on noncontributory social security programs; in contrast, higher levels of formality increase the government's SSCs, leaving total expenditures virtually unchanged. Overall, these changes to the SUFE would increase the fiscal balance from 6.87 to 7.49 percent of GDP while substantially raising the formality rate.

Simulation of Changes to the PIT Scheme

Table 5 presents the impact of simulated changes to the PIT scheme. Panel A shows the baseline, which reflects the conditions that were in place in 2018. The first reform scenario grants a 100 percent tax exemption to formal workers in the first two income brackets. This exemption is applied by setting a 0 percent tax rate and a fixed payment amount of MXN 0 for the first two brackets, while all other tax rates remain unchanged. To avoid creating a tax notch, the fixed amount for the third income bracket would also be MXN 0, and the fixed amounts for the remaining brackets would be adjusted according to the formula currently used by the Ministry of Finance. Panel B of table 5 shows the PIT table for the proposed reform after combining the first two income brackets in panel A. As described next, eliminating PIT for workers in the lowest income brackets could significantly increase employment formality with almost no effect on the fiscal balance.

Table 6 shows the impact of these changes to the PIT scheme. Exempting incomes of up to MXN 4,910 per month from PIT liability increases the net incomes of low-wage formal workers, encouraging high-skilled informal workers to formalize. As a result, the formality rate increases by almost ten percentage points. Formalization among high-skilled informal workers reduces the average net earnings of both formal and informal employees as workers who were previously the highest-paid employees in the informal sector become the lowest-paid employees in the formal sector, while the inflow of low-skill workers into the formal sector further reduces its equilibrium wage rate. Although the outflow of labor from the informal sector raises its equilibrium wage rate, this effect is more than offset by the exit of highly paid workers from the informal sector, combined with an influx of formerly own-account workers seeking higher incomes as informal wage employees.

These changes to the PIT entail multiple countervailing effects on the fiscal accounts. Eliminating the tax liability of low-income workers causes

TABLE 6. Personal Income Tax (PIT) Reform Simulations

Variable	*Data*	*Baseline calibration*	*Income tax exemption up to MXN 4,910*
Occupation (as a share of employment)			
Total informal	0.658	0.658	0.559
Own-account	0.263	0.263	0.183
Informal wage	0.395	0.395	0.376
Formal wage	0.294	0.294	0.393
Employers	0.048	0.048	0.048
Average net income (pesos per month)			
Total wage	5,668	5,720	5,549
Formal	7,447	7,469	6,978
Informal	4,344	4,419	4,054
Own-account	4,762	2,766	2,547
Employers	12,817	12,817	13,376
Fiscal accounts (% of GDP)			
Wage workers			
(A) Income tax	3.10	3.10	2.67
(B) SSC	n.a.	0.29	0.36
Employers			
(C) Income tax	3.10	3.10	2.80
(D) SSC	n.a.	2.81	3.51
Government: Contributory SS			
(E) SS revenue (B + D)	3.10	3.10	3.87
(F) SS expenditures	0.53	0.53	0.67
(G) Extra operating expenditures	. . .	. . .	0.24
(H) Balance (E − F − G)	2.57	2.57	2.96
Government: Other			
(I) Income tax revenues (A + C)	6.20	6.20	5.47
(J) Subsidy to formal employment	0.20	0.20	0.27
(K) Noncontributory SS	1.70	1.70	1.38
(L) Balance (I − J − K)	4.30	4.30	3.82
Government: Total			
(M) Revenue (E + I)	9.30	9.30	9.34
(N) Expenditures (F + G + J + K)	2.43	2.43	2.56
(O) Budget balance (M − N)	6.87	6.87	6.78
Tax burden (%)			
Income tax			
Wage workers	50.0	50.0	48.8
Employers	50.0	50.0	51.2
SSC			
Wage workers	n.a.	9.4	9.3
Employers	n.a.	90.6	90.7

Source: Authors' elaboration.
. . . Not applicable.
n.a. Not available.

PIT revenue to fall to 2.67 percent of GDP, while the rising formalization rate allows firms to increase their tax deductions, causing a slight drop in CIT revenue. Since the former effect is greater than the latter, the relative income tax burden on firms rises from 50 to 51.2 percent. While income tax revenue declines, formalization increases SSCs, leaving fiscal revenue largely unchanged. Meanwhile, total expenditures increase from 2.43 to 2.56 percent of GDP because of higher spending on SSC, the SUFE, and health care (row G, extra operating expenditures). Consequently, the budget balance deteriorates slightly relative to the baseline.

Simulation of Simultaneous Changes to the PIT and SUFE

This section simulates the effects of modifying both the PIT and SUFE schemes, a scenario described in the tables as the full reform. The change to the PIT is the same as that described in the previous section, while the change to the SUFE is the scenario in which a unform transfer of MXN 400 per month is granted to all workers with an income of up to MXN 4,910 per month, with transfer amounts being progressively reduced above that level and ultimately eliminated for earnings of MXN 7,410 per month or more. As the changes to the PIT and SUFE both incentivize formalization individually, their combined effect is especially large. Under the full reform scenario, the formality rate rises by nearly 12 percentage points. Meanwhile, the positive fiscal impact of the change in the SUFE outweighs the negative impact of the PIT change, resulting in a modest net improvement in the government's budget balance.

Table 7 presents the model's results in terms of employment, average earnings, and the average utility of disposable income. The combination of the PIT exemption for the first two income brackets and the redesign of the SUFE increases the share of formal wage employees from 29.4 to 41.3 percent of total employment. As discussed above, these measures strongly incentivize labor formalization mainly at the expense of own-account workers. They also decrease average net earnings for both formal and informal employment as the most highly skilled own-account and informal wage workers enter informal wage and formal wage occupations, respectively.

Table 7 also shows the average utility of disposable income for workers in both scenarios, monetized in pesos per month. In our model, workers' utility is equivalent to their earnings, given by equations 9, 10, and 11. Therefore, the difference between the average utility of disposable income and net income for both own-account workers and informal wage employees is explained by their valuation of lump sum transfers. For formal employees, the difference

TABLE 7. Effects of the Combined PIT and SUFE Reforms on Occupational Choice, Net Income, and Average Utility of Disposable Income

Variable	*Data*	*Baseline model*	*Full reform model*
Occupation (as a share of employment)			
Total informal	0.658	0.658	0.539
Own-account	0.263	0.263	0.179
Informal wage	0.395	0.395	0.360
Formal wage	0.294	0.294	0.413
Employers	0.048	0.048	0.048
Average net income (pesos per month)			
Total wage	5,668	5,720	5,545
Formal	7,447	7,469	6,700
Informal	4,344	4,419	4,223
Own-account	4,762	2,766	2,536
Employers	12,817	12,817	13,331
Average utility of disposable income (pesos per month)			
Formal wage	. . .	8,441	7,556
Informal wage	. . .	5,225	5,029
Own-account	. . .	3,572	3,342
Employers	. . .	12,817	13,331

Source: Authors' elaboration.
. . . Not applicable.

reflects the valuation of their social security benefits plus the fringe benefits conferred by formal employment, and thus the utility and net income of these workers move in the same direction. For entrepreneurs, utility is assumed to be identical to net benefits.

Table 8 shows how the combined changes to the PIT and SUFE affect the fiscal accounts. Under this scenario, PIT revenue increases slightly relative to the scenario in which the PIT is reformed while the SUFE is left unchanged (see table 6). This effect occurs for two reasons. First, because the modified SUFE is less favorable for wage earners with incomes greater than MXN 5,600 per month, the SUFE reform increases the amount of PIT collected from workers with incomes between MXN 5,600 and MXN 7,400 per month. Second, higher formalization expands the PIT tax base. This increase in formalization also raises revenue from SSCs. Overall, the changes implemented under the full reform scenario increase total government revenue by 0.34 percent of GDP relative to the baseline.

Table 8 also shows the effects of the combined PIT and SUFE reforms on public spending. Higher formality rates increase expenditures on contributory social security to 0.71 percent of GDP while boosting extra operating expenses

TABLE 8. Effects of the Combined PIT and SUFE Reforms on the Fiscal Accounts

Variable	*Data*	*Baseline model*	*Full reform model*
Fiscal accounts (% of GDP)			
Wage workers			
(A) Income tax	3.10	3.10	2.90
(B) SSC	n.a.	0.29	0.38
Employers			
(C) Income tax	3.10	3.10	2.81
(D) SSC	n.a.	2.81	3.70
Government: Contributory SS			
(E) SS revenue (B + D)	3.10	3.10	4.08
(F) SS expenditures	0.53	0.53	0.71
(G) Extra operating expenditures	...	...	0.27
(H) Balance (E − F − G)	2.57	2.57	3.09
Government: Other			
(I) Income tax revenues (A + C)	6.20	6.20	5.71
(J) Subsidy to formal employment	0.20	0.20	0.25
(K) Noncontributory SS	1.70	1.70	1.34
(L) Balance (I − J − K)	4.30	4.30	4.12
Government: Total			
(M) Revenue (E + I)	9.30	9.30	9.78
(N) Expenditures (F + G + J + K)	2.43	2.43	2.57
(O) Budget balance (M − N)	6.87	6.87	7.21
Tax burden (%)			
Income tax			
Wage workers	50.0	50.0	50.9
Employers	50.0	50.0	49.2
SSC			
Wage workers	n.a.	9.4	9.3
Employers	n.a.	90.6	90.7

Source: Authors' elaboration.
... Not applicable.
n.a. Not available.

in health care by 0.27 percent of GDP. However, formalization also decreases spending on noncontributory social security transfers to informal workers, and total government spending increases by just 0.14 percent of GDP. Because the increase in revenues exceeds the increase in expenditures, the budget balance improves relative to the baseline. The combined reforms successfully encourage labor formalization while also strengthening the fiscal accounts, yielding clear benefits in two major economic policy areas while incurring no evident cost.

Finally, table 8 reports how the burden of income taxes and SSCs is distributed between workers and employers. The reform slightly increases the income tax burden borne by workers from 50.0 to 50.9 percent, but the SSC burden remains broadly unchanged.

Sensitivity Analysis

The results presented above are based on specific parameter values. However, some values are determined a priori, as no comparable evidence is presented in the literature. To test the robustness of the results, we conduct a sensitivity analysis of the full reform scenario. The analysis shows that the combined PIT and SUFE reforms generate a significant increase in employment formality and a modest improvement in the fiscal accounts under a range of alternative parameter values.

The following parameters are considered for this analysis: the parameter related to the elasticity of substitution between formal and informal wage labor (ψ); returns to scale for the production function of the entrepreneur (γ); the valuation of social security benefits (β_F) and noncontributory transfers (β_I); returns to scale for the skill levels of informal wage workers (α); and the lower and upper bounds for labor ability ($\underline{e}$ and $\overline{e}$). The first two parameters are directly related to the labor demand of entrepreneurs; the next three are labor supply parameters affecting the occupational choices of workers; and the last two relate to the distribution of labor ability. In each of the following exercises, a single parameter is changed, and the ten parameters under the third group are reestimated to match the moments of table 2. We also present an exercise in which three parameters are simultaneously changed.

For ψ, we consider the alternative values of $\psi = 0$ and $\psi = -9$, which imply elasticities of substitution between formal and informal labor of 1 and 0.1, respectively, instead of the elasticity value of 10 used in the baseline scenario. Lower values for ψ reflect a diminished willingness among employers to substitute formal for informal labor, which attenuates the impact of the reforms on formalization. For parameter γ, alternative values of 0.67 and 0.82 are adopted instead of the original value of 0.76. A lower value for γ implies a decrease in the marginal product of labor, which discourages the hiring of wage workers, while a higher value implies the opposite. Decreasing the value of parameter β_F from 0.30 to 0.05 means that workers value their social security benefits 83 percent less than in the baseline scenario, while increasing the value of β_F from 0.30 to 0.60 means that they value those benefits twice as highly as in the baseline. Raising the value for β_I from 0.85 to 1 increases the valuation of lump sum transfers, encouraging informal employment at the expense of formal employment, whereas reducing the value for β_I to 0.35 has the opposite effect. For α, we use an alternative value of 0.9, which is slightly above the benchmark value of 0.857. For labor ability, we raise the lower bound of e while lowering its upper bound, narrowing the domain of the distribution, which has unpredictable implications for the impact of the reforms on the formalization rate.

Tables 9, 10, and 11 present the results of the sensitivity analysis, as well as the simulation conducted under the baseline scenario. We begin by analyzing changes in a single parameter. For occupational choices, the strong effect on formality reported above is robust to alternative parameter values. Even in the least favorable scenario ($\psi = 0$), the formality rate increases to 37 percent. Changes in average net income relative to the baseline are registered across all occupation types, but these changes are modest. Changes to the fiscal accounts are also relatively small but uniformly positive: in the least favorable scenario, the budget balance rises from 6.87 percent of GDP to 7.05 percent.

The last column of table 11 shows the results of a simulation in which three parameters are simultaneously changed to make formalization more difficult. The elasticity of substitution between formal and informal wage workers is set to 1 ($\psi = 0$), while β_F and β_I are set at 0.05 and 1, respectively. Even in this scenario, the formality rate rises by seven percentage points over the baseline, from 29.4 to 36.4 percent, while the fiscal balance remains broadly unchanged.

The sensitivity analysis indicates that the results obtained are robust to a range of alternative parameter values. Even under the least favorable scenario, the combined reforms would have a highly positive impact on formalization while incurring no significant fiscal cost.

Conclusion

This paper has presented a static general equilibrium model of occupational choice with heterogeneous labor and entrepreneurial skills to evaluate how changes in the labor income tax scheme would affect employment informality and the fiscal accounts. Heterogeneity in labor skills is important because it generates an income distribution and a corresponding income-based tax and subsidy structure, capturing an important characteristic of tax schemes observed in countries around the world. Heterogeneity in entrepreneurial skills is also relevant because it allows larger firms to hire more formal workers than smaller

TABLE 9. Sensitivity Analysis of the Combined PIT and SUFE Reforms: Labor Demand Parameters

Variable	*Benchmark model: full reform*	*Sensitivity analysis*			
		$\psi = 0$	$\psi = -9$	$\gamma = 0.67$	$\gamma = 0.82$
Occupation (as a share of employment)					
Total informal	0.539	0.582	0.581	0.558	0.529
Own-account	0.179	0.179	0.171	0.195	0.162
Informal wage	0.360	0.403	0.410	0.362	0.367
Formal wage	0.413	0.370	0.371	0.394	0.423
Employers	0.048	0.048	0.048	0.048	0.048
Average net income (pesos per month)					
Total wage	5,545	5,539	5,521	5,579	5,509
Formal	6,700	6,742	6,748	6,707	6,702
Informal	4,223	4,433	4,411	4,351	4,136
Own-account	2,536	2,536	2,510	2,583	2,482
Employers	13,331	13,374	13,383	13,214	13,377
Fiscal accounts (% of GDP)					
Wage workers					
(A) Income tax	2.90	2.82	2.83	2.79	3.03
(B) SSC	0.38	0.35	0.35	0.37	0.39
Employers					
(C) Income tax	2.81	3.07	3.07	2.93	2.73
(D) SSC	3.70	3.26	3.27	3.55	3.79
Government: Contributory SS					
(E) SS revenue (B + D)	4.08	3.61	3.62	3.92	4.18
(F) SS expenditures	0.71	0.61	0.61	0.68	0.72
(G) Extra operating expenditures	0.27	0.21	0.21	0.23	0.30
(H) Balance (E − F − G)	3.09	2.79	2.80	3.00	3.16
Government: Other					
(I) Income tax revenues (A + C)	5.71	5.89	5.89	5.72	5.76
(J) Subsidy to formal employment	0.25	0.19	0.20	0.24	0.25
(K) Noncontributory SS	1.34	1.44	1.44	1.40	1.31
(L) Balance (I − J − K)	4.12	4.26	4.25	4.08	4.20
Government: Total					
(M) Revenue (E + I)	9.78	9.50	9.51	9.64	9.93
(N) Expenditures (F + G + J + K)	2.57	2.45	2.46	2.55	2.58
(O) Budget balance (M − N)	7.21	7.05	7.05	7.08	7.35
Tax burden (%)					
Income tax					
Wage workers	50.9	47.9	48.0	48.8	52.6
Employers	49.2	52.1	52.0	51.2	47.4
SSC					
Wage workers	9.3	9.6	9.7	9.3	9.3
Employers	90.7	90.4	90.3	90.7	90.7

Source: Authors' elaboration.

TABLE 10. Sensitivity Analysis of the Combined PIT and SUFE Reforms: Labor Supply Parameters

Variable	*Benchmark model: full reform*	*Sensitivity analysis*				
		$\beta_F = 0.05$	$\beta_F = 0.60$	$\beta_I = 0.35$	$\beta_I = 1$	$\alpha = 0.9$
Occupation (as a share of employment)						
Total informal	0.539	0.543	0.549	0.510	0.548	0.529
Own-account	0.179	0.185	0.189	0.194	0.173	0.205
Informal wage	0.360	0.357	0.360	0.316	0.375	0.323
Formal wage	0.413	0.409	0.403	0.442	0.404	0.423
Employers	0.048	0.048	0.048	0.048	0.048	0.048
Average net income (pesos per month)						
Total wage	5,545	5,515	5,486	5,304	5,481	5,705
Formal	6,700	6,743	6,568	5,995	6,790	6,446
Informal	4,223	4,108	4,276	4,338	4,070	4,735
Own-account	2,536	2,640	2,249	1,876	2,547	2,414
Employers	13,331	13,293	13,320	13,289	13,315	13,184
Fiscal accounts (% of GDP)						
Wage workers						
(A) Income tax	2.90	2.93	2.72	3.16	2.87	3.02
(B) SSC	0.38	0.38	0.37	0.41	0.37	0.40
Employers						
(C) Income tax	2.81	2.80	2.87	2.86	2.81	2.84
(D) SSC	3.70	3.65	3.66	4.00	3.61	3.85
Government: Contributory SS						
(E) SS revenue (B + D)	4.08	4.02	4.03	4.42	3.99	4.25
(F) SS expenditures	0.71	0.70	0.71	0.77	0.69	0.74
(G) Extra operating expenditures	0.27	0.27	0.24	0.33	0.26	0.29
(H) Balance (E − F − G)	3.09	3.05	3.08	3.32	3.04	3.22
Government: Other						
(I) Income tax revenues (A + C)	5.71	5.72	5.59	6.01	5.68	5.86
(J) Subsidy to formal employment	0.25	0.25	0.25	0.26	0.24	0.26
(K) Noncontributory SS	1.34	1.35	1.37	1.27	1.36	1.33
(L) Balance (I − J − K)	4.12	4.13	3.97	4.48	4.08	4.27
Government: Total						
(M) Revenue (E + I)	9.78	9.74	9.61	10.43	9.67	10.11
(N) Expenditures (F + G + J + K)	2.57	2.56	2.57	2.63	2.55	2.61
(O) Budget balance (M − N)	7.21	7.18	7.05	7.80	7.11	7.50
Tax burden (%)						
Income tax						
Wage workers	50.9	51.2	48.6	52.5	50.5	51.6
Employers	49.2	48.8	51.4	47.5	49.5	48.4
SSC						
Wage workers	9.3	9.3	9.3	9.3	9.3	9.3
Employers	90.7	90.7	90.7	90.7	90.7	90.7

Source: Authors' elaboration.

TABLE 11. Sensitivity Analysis of the Combined PIT and SUFE Reforms: Skills Distribution and Multiple Changes

Variable	*Benchmark model: full reform*	*Sensitivity analysis*	
		$\underline{e} = 0.3, \overline{e} = 10$	$\psi = 0, \beta_F = 0.01, \beta_I = 1$
Occupation (as a share of employment)			
Total informal	0.539	0.531	0.588
Own-account	0.179	0.182	0.171
Informal wage	0.360	0.349	0.417
Formal wage	0.413	0.421	0.364
Employers	0.048	0.048	0.048
Average net income (pesos per month)			
Total wage	5,545	5,537	5,454
Formal	6,700	6,560	6,948
Informal	4,223	4,301	4,151
Own-account	2,536	2,722	2,618
Employers	13,331	13,336	13,482
Fiscal accounts (% of GDP)			
Wage workers			
(A) Income tax	2.90	2.76	2.74
(B) SSC	0.38	0.38	0.34
Employers			
(C) Income tax	2.81	2.80	3.05
(D) SSC	3.70	3.79	3.19
Government: Contributory SS			
(E) SS revenue (B + D)	4.08	4.17	3.53
(F) SS expenditures	0.71	0.74	0.60
(G) Extra operating expenditures	0.27	0.28	0.19
(H) Balance (E − F − G)	3.09	3.15	2.74
Government: Other			
(I) Income tax revenues (A + C)	5.71	5.56	5.79
(J) Subsidy to formal employment	0.25	0.27	0.20
(K) Noncontributory SS	1.34	1.32	1.45
(L) Balance (I − J − K)	4.12	3.97	4.15
Government: Total			
(M) Revenue (E + I)	9.78	9.72	9.33
(N) Expenditures (F + G + J + K)	2.57	2.60	2.44
(O) Budget balance (M − N)	7.21	7.12	6.88
Tax burden (%)			
Income tax			
Wage workers	50.9	49.6	47.4
Employers	49.2	50.4	52.6
SSC			
Wage workers	9.3	9.2	9.7
Employers	90.7	90.8	90.3

Source: Authors' elaboration.

firms, reflecting another important feature of the data. The model has been calibrated for Mexico, which is characterized by a high rate of labor informality. The analysis included various reforms to the current SUFE and PIT schemes, both separately and together, and their estimated impacts on labor formality and the fiscal accounts.

The exercises indicate that minor modifications to the current labor tax and subsidy scheme could have large positive effects on labor formality with no adverse impact on the fiscal balance. Modifying the SUFE schedule while eliminating the PIT liability of the lowest-income formal workers strongly incentivizes formalization, and our simulations suggest that these measures could increase the formality rate by between 7.0 and 11.9 percentage points. Importantly, these changes to tax and subsidy policies would entail no net cost to the government: the fiscal balance would either remain constant or modestly improve. Meanwhile, the distribution of the income tax burden between workers and employers would shift only slightly relative to the baseline scenario.

The simulation exercises presented above underscore how general equilibrium models can yield important insights into prospective changes to income tax and subsidy policies in contexts of high labor informality. For example, in the scenario where the SUFE is eliminated, the government balance does not improve in response to the decrease in subsidy spending, but rather deteriorates because of a sharp increase in informality. Similarly, reducing the PIT liability for low-income workers can improve the fiscal balances by increasing the formality rate, which more than compensates for the loss of direct tax revenue. These counterintuitive results are better understood once the endogenous links between informality and the tax base have been considered.

Despite the important results obtained by the simulations, the model could be developed further to address some of its limitations. For example, because of its static nature, the model is unable to produce a transitional path to a new equilibrium after the introduction of fiscal changes. During the transitional period, such changes may be substantially different from those observed under the new equilibrium. Another extension relates to the inclusion of informal firms that hire wage workers, which could enable the model to explicitly evaluate the effect of changes in the labor income tax scheme on those firms.

References

Acemoglu, Daron, and David Autor. 2011. “Skills, Tasks, and Technologies: Implications for Employment and Earnings.” In *Handbook of Labor Economics,* vol. 4B, ed. Orley Ashenfelter and David Card, pp. 1043–71. Amsterdam: Elsevier.

Albrecht, James, Lucas Navarro, and Susan Vroman. 2009. “The Effects of Labour Market Policies in an Economy with an Informal Sector.” *Economic Journal* 119 (539): 1105–29.

Allub, Lian, and Andrés Erosa. 2019. “Financial Frictions, Occupational Choice, and Economic Inequality.” *Journal of Monetary Economics* 107 (November): 63–76.

Antón, Arturo, and Fausto Hernández. 2017. “Reforma de la hacienda pública y de la seguridad social en México.” Working Paper 005/2017. Mexico City: Centro de Estudios Espinosa Yglesias.

Antón, Arturo, Fausto Hernández, and Santiago Levy. 2012. *The End of Informality in Mexico? Fiscal Reform for Universal Social Insurance*. Washington, DC: Inter-American Development Bank.

Bobba, Matteo, Luca Flabbi, and Santiago Levy. 2022. “Labor Market Search, Informality, and Schooling Investments.” *International Economic Review* (forthcoming).

Bosch, Mariano, and Julen Esteban-Pretel. 2012. “Job Creation and Job Destruction in the Presence of Informal Markets.” *Journal of Development Economics* 98 (2): 270–86.

———. 2015. “The Labor Market Effects of Introducing Unemployment Benefits in an Economy with High Informality.” *European Economic Review* 75 (April): 1–17.

Dagum, Camilo. 1977. “A New Model of Personal Income Distribution: Specification and Estimation.” *Économie Appliquée* 30 (3): 413–37.

de Paula, Aureo, and Jose A. Scheinkman. 2010. “Value-Added Taxes, Chain Effects, and Informality.” *American Economic Journal: Macroeconomics* 2 (4): 195–221.

D’Erasmo, Pablo N., Hernan J. Moscoso, and Asli Senkal. 2014. “Misallocation, Informality, and Human Capital: Understanding the Role of Institutions.” *Journal of Economic Dynamics and Control* 42 (May): 122–42.

Fortin, Bernard, Nicholas Marceau, and Luc Savard. 1997. “Taxation, Wage Controls, and the Informal Sector.” *Journal of Public Economics* 66 (2): 293–312.

Galiani, Sebastian, and Federico Weinschelbaum. 2012. “Modeling Informality Formally: Households and Firms.” *Economic Inquiry* 50 (3): 821–38.

Gollin, Douglas. 2008. “Nobody’s Business but My Own: Self-employment and Small Enterprise in Economic Development.” *Journal of Monetary Economics* 55 (2): 219–33.

Gong, Xiaodong, and Arthur van Soest. 2002. “Wage Differentials and Mobility in the Urban Labour Market: A Panel Data Analysis for Mexico.” *Labour Economics* 9 (4): 513–29.

Günther, Isabel, and Andrey Launov. 2012. “Informal Employment in Developing Countries: Opportunity or Last Resort?” *Journal of Development Economics* 97 (1): 88–98.

ILO (International Labour Organization). 2018. *Women and Men in the Informal Economy: A Statistical Picture*, 3rd ed. Geneva: ILO.

IMSS (Mexican Institute of Social Security). 2018. *Informe al Ejecutivo Federal y al Congreso de la unión sobre la situación financiera y los riesgos del Instituto Mexicano del Seguro Social 2017–2018*. Mexico City: IMSS.

INEGI (National Institute of Statistics and Geography). 2013. *Encuesta Nacional de Ocupación y Empleo.* Third quarter. Mexico City: INEGI.

———. 2014. *Censos Económicos 2014*. Mexico City: INEGI.

———. 2017. *Sistema de Cuentas Nacionales de México*. Mexico City: INEGI.

Jovanovic, Boyan. 1994. "Firm Formation with Heterogeneous Management and Labor Skills." *Small Business Economics* 6 (3): 185–91.

Leal, Julio. 2014. "Tax Collection, the Informal Sector, and Productivity." *Review of Economic Dynamics* 17 (2): 262–86.

Levy, Santiago. 2008. *Good Intentions, Bad Outcomes: Social Policy, Informality, and Economic Growth in Mexico*. Washington, DC: Brookings Institution Press.

López, José J. 2017. "A Quantitative Theory of Tax Evasion." *Journal of Macroeconomics* 53 (September):107–26.

Lucas, Robert E. 1978. "On the Size Distribution of Business Firms." *Bell Journal of Economics* 9 (2): 508–23.

Margolis, David N., Lucas Navarro, and David A. Robalino. 2014. "Unemployment Insurance, Job Search, and Informal Employment." In *Social Insurance, Informality, and Labor Markets*, ed. Markus Frölich and others. Oxford: Oxford University Press.

McDonald, James B., and Michael Ransom. 2008. "The Generalized Beta Distribution as a Model for the Distribution of Income: Estimation of Related Measures of Inequality." In *Modeling Income Distributions and Lorenz Curves*, ed. Duangkamon Chotikapanich. New York: Springer.

Meghir, Costas, Renata Narita, and Jean-Marc Robin. 2015. "Wages and Informality in Developing Countries." *American Economic Review* 105 (4): 1509–46.

Ministry of Finance of Mexico. 2017a. *Cuenta Pública 2017*. Mexico City.

———. 2017b. *Presupuesto de Gastos Fiscales 2017*. Mexico City.

———. 2017c. *Informe sobre la Situación Económica, las Finanzas Públicas y la Deuda Pública 2017*. Fourth quarter. Mexico City.

Narita, Renata. 2020. "Self-employment in Developing Countries: A Search Equilibrium Approach." *Review of Economic Dynamics* 35 (January): 1–34.

OECD/IDB/CIAT (Organization for Economic Cooperation and Development, Inter-American Development Bank, and Inter-American Center of Tax Administrations). 2016. *Taxing Wages in Latin America and the Caribbean*. Paris: OECD Publishing.

Perry, Guillermo, William F. Maloney, Omar S. Arias, and others. 2007. *Informality: Exit and Exclusion*. Washington, DC: World Bank.

Poschke, Markus. 2013. "Who Becomes an Entrepreneur? Labor Market Prospects and Occupational Choice." *Journal of Economic Dynamics and Control* 37 (3): 693–710.

Pratap, Sangeeta, and Erwan Quintin. 2008. "The Informal Sector in Developing Countries: Output, Assets, and Employment." In *Personal Wealth from a Global Perspective*, ed. James B. Davies. Oxford: Oxford University Press.

Rauch, James E. 1991. "Modelling the Informal Sector Formally." *Journal of Development Economics* 35 (1): 33–47.

Samaniego de la Parra, Brenda. 2017. "Formal Firms, Informal Workers, and Household Labor Supply in Mexico." Ph.D. diss., University of Chicago.

Saracoglu, Durdane S. 2008. "The Informal Sector and Tax on Employment: A Dynamic General Equilibrium Investigation." *Journal of Economic Dynamics and Control* 32 (2): 529–49.

Summers, Lawrence H. 1989. "Some Simple Economics of Mandated Benefits." *American Economic Review: Papers and Proceedings* 79 (2): 177–83.

Ulyssea, Gabriel. 2010. "Regulation of Entry, Labor Market Institutions, and the Informal Sector." *Journal of Development Economics* 91 (1): 87–99.

———. 2018. "Firms, Informality, and Development: Theory and Evidence from Brazil." *American Economic Review* 108 (8): 2015–47.

Zenou, Yves. 2008. "Job Search and Mobility in Developing Countries: Theory and Policy Implications." *Journal of Development Economics* 86 (2): 336–55.

ALI ALICHI
International Monetary Fund, United States

IPPEI SHIBATA
International Monetary Fund, United States

KADIR TANYERI
International Monetary Fund, United States

Fiscal Policy Multipliers in Small States

ABSTRACT This paper estimates fiscal policy multipliers for small states using two distinct models: an empirical forecast error model with data from twenty-three small states across the world, and a dynamic stochastic general equilibrium (DSGE) model calibrated to a hypothetical small state's economy. We find that, in the short term, multipliers for government consumption and investment in small states are both about 0.4, on average, for empirical and DSGE baseline results, and they are affected by imports as a share of GDP, the level of government debt, and the economy's position in the business cycle, among other factors. In the medium to long run, while fiscal policy using government consumption is ineffective, government investment has a multiplier of about 0.7, on average, for empirical and DSGE baseline results. These results are robust to different model specifications and characteristics of small states. Inability to affect GDP using government consumption could be frustrating for policymakers when an expansionary policy is needed but encouraging when they consider fiscal consolidation.

JEL Codes: E62, C3

Keywords: Government spending, fiscal policy, fiscal multipliers, small states

The International Monetary Fund (IMF) defines thirty-four developing member countries with populations of fewer than 1.5 million to be small states.[1] These small states are spread across Africa, Asia, the Caribbean, and Europe. Small states are characterized by limited economic scale, including small populations, narrow production bases, and limited opportunities for diversification. A small population base implies low demand for services and

ACKNOWLEDGMENTS We thank the editor, Bernardo Guimarães, and Philip Barrett, Takuji Komatsuzaki, Hiroaki Miyamoto, and Machiko Narita for their invaluable suggestions, and Lulu Shui and Heidi Canelas for excellent research assistance and formatting, respectively. The views expressed in this paper are those of the authors and do not necessarily represent the views of the IMF, its Executive Board, or IMF management.

1. The definition of small states in this paper follows IMF (2017). This definition differs slightly from the World Bank's definition of small states, which includes fifty countries that have a population of 1.5 million or fewer or are members of the Small States Forum—a high-level meeting of policymakers hosted by the World Bank during the IMF-World Bank Annual Meetings (World Bank, 2016). For more information, see World Bank in Small States: Overview (www.worldbank.org/en/country/smallstates/overview).

limited interest from international investors in the country. Some small states struggle with geographic remoteness and are also more prone to experience the effects of climate change and natural disasters.

Many small states have had large overall fiscal deficits over the past three decades. As a percentage of gross domestic product (GDP), average government expenditures in these countries have been increasing over time, diverging from government tax revenues, which have increased only slightly (IMF, 2018b). Following the global financial crisis of 2008–09, government deficits in small states have remained high in response to various exogenous shocks, including commodity price increases, natural disasters, and exchange rate depreciations. The increase in government spending has been mostly in current government spending, while capital spending has remained modest (IMF, 2018b). Low investment content would have lasting negative effects on the economy.

For many small states, fiscal consolidation is necessary to put public finances on a sustainable path and open fiscal space to confront future adverse economic shocks. However, the first question that policymakers usually ask when considering fiscal consolidation is how it would affect GDP growth. This paper provides an answer to this question by estimating fiscal policy multipliers—the impact of fiscal policy on GDP—for small states.

The high import share of GDP in small states points to possibly lower fiscal multipliers than in the case of larger developing and advanced economies. This could be because small states are generally more open than larger economies. In a standard textbook Mundell-Fleming model, the fiscal multiplier in a more open economy would be lower because part of the increase in aggregate demand that is boosted by fiscal policy would be spent on imports, reducing net exports. Empirical evidence on larger countries supports this conjecture. Ilzetzki, Mendoza, and Végh (2013) show that fiscal multipliers in open economies are indeed smaller than in closed economies.

Higher government debt levels of small states also point to possibly lower multipliers than other countries.[2] While the literature is very thin on small states, Ilzetzki, Mendoza, and Végh (2013) show for larger countries that fiscal

2. Average gross (external) debt for small states in 2018 was 58.7 (53.9) percent of GDP, whereas larger developing economies had average gross (external) debt of 53.2 (47.8) percent. There are several reasons why small states might have a high level of debt, including their exposure to natural disasters, vulnerability to external shocks, and shallow financial systems.

multipliers in high-debt economies are negative (while those of low-debt economies are positive). This is because for countries with high government debt levels, an increase in government spending could signal fiscal tightening in the near future and dampen the fiscal multiplier impact.

The main contribution of this paper is to estimate fiscal policy multipliers in small states using two distinct models: an empirical model, which we argue is more reliable than prior ones, and an open economy dynamic stochastic general equilibrium (DSGE) model. The empirical model uses the forecast error and a local projection method, as in Jordà (2005), to estimate the causal impact of a change in government consumption or government investment on GDP—namely, fiscal multipliers.[3] The DSGE model in this paper is the IMF global integrated monetary and fiscal (GIMF) model, which we calibrate to a hypothetical small open economy. The results suggest that in the short term, which we define as two years from the initial shock (with 0 being the initial impact year), the government consumption multiplier (on the level of GDP) is about 0.3 using our empirical model and about 0.6 using our GIMF model (average of 0.4, over the two models). Short-term government investment multipliers are estimated at 0.1 using our empirical model, and 0.7 using our GIMF model (average of 0.4). In the medium term (defined in this paper as three to four years), the government consumption multipliers are around zero, while the government investment multipliers are around 0.7, on average, in both the empirical and GIMF models. Sensitivity analysis shows that the GIMF multipliers could be smaller or larger depending on many factors, including imports as a share of GDP, the level of government debt, and the economy's position in the business cycle.

3. Owing to a lack of data, we are not able to separate government consumption and transfers for most small states in our sample. Therefore, in our empirical work (only), we use the concept of government current primary spending, which is total government spending minus investment and interest expenses—in other words, government consumption plus transfers. In our GIMF simulations, we separate government consumption and transfers. We do not estimate tax multipliers in our empirical model, since tax revenues are known to be highly endogenous to the conditions of the economy, and even our forecast error methods cannot account for the endogeneity issues in tax revenues (Furceri and others, 2018). This is because tax reforms and other discretionary changes to taxes are very infrequent in all countries. In most years, most of the changes in tax/GDP are due to automatic stabilizers built into the tax systems. On the other hand, discretionary fiscal spending decisions change almost every year in almost all countries, for example, in the context of annual budget processes. Moreover, the portion of spending that is affected by automatic stabilizers (such as the social protection portion) is generally a small part of the total spending.

Our findings of small government consumption multipliers and relatively large government investment multipliers in small states in the medium term are both in line with the existing literature. For example, Gonzalez-Garcia, Lemus, and Mrkaic (2013), Guy and Belgrave (2012), and Narita (2014) find similar results for a group of Caribbean countries using a structural vector autoregression (SVAR) and a dynamic panel framework. There is also a vast literature on estimating fiscal multipliers for larger countries. However, while there are many similarities in terms of methodology, the results of this strand of literature are of limited use for small states, given the aforementioned characteristics of these countries.

Our empirical approach has many advantages over previous studies for small states, making our multiplier estimates more reliable. Relative to the previous studies, this paper (1) has a larger sample size; (2) estimates flexible nonlinear fiscal multipliers via a local projection method; (3) does not rely on interpolating quarterly data from annual series; (4) mitigates, via a forecast error approach, the anticipation or foresight problem, in which agents change their behavior in anticipation of future changes in fiscal variables; (5) estimates state-dependent fiscal multipliers (expansion versus consolidation and boom versus recession); and (6) estimates fiscal multipliers for government consumption and investment separately (see our literature review below and the detailed explanation of the advantages of our empirical approach).[4]

The rest of the paper is organized as follows. The next section reviews the empirical model employed in this study. Subsequently we describe the data and present empirical results. After describing the GIMF model used in this study, we present the results from the GIMF model and compare the results with other studies. The final section concludes.

Empirical Model: The Forecast Error Approach

Our empirical model uses the forecast error approach. The idea behind this approach is that the forecast captures agents' anticipation of fiscal actions, and the deviation of reality from that forecast—that is, the forecast error—plausibly captures an unanticipated increase or decrease in government spending.

4. This anticipation or foresight problem could be severe when estimating fiscal multipliers using annual rather than quarterly data.

Identification of Fiscal Shocks Using WEO Vintage Data

For the fiscal variables and shocks, we use vintage data from past issues of the IMF's October publication of the *World Economic Outlook* (WEO), following Furceri and Li (2017). In the WEO, macroeconomic variables are reported at an annual frequency for IMF member countries, and forecasts are made by IMF staff for the projection years. In the October WEO, forecasts for that year are made based on all the information that is available to the IMF country teams. Forecast errors are constructed from government consumption and government investment, as a percentage of GDP. We calculate the unanticipated fiscal variable shock, $\text{FShock}^k_{j,t}$, as the difference between the actual and forecast fiscal variables:

$$\text{FShock}^k_{j,t} = f^{k,Actual}_{j,t} - f^{k,Forecast}_{j,t}, \tag{1}$$

where $f^k_{j,t} \equiv F^k_{j,t}/Y_{j,t-1}$ is a fiscal variable ($F^k_{j,t}$) of type $k \in \{I, C\}$ as a percentage of the previous year's GDP ($Y_{j,t-1}$). The fiscal variable is either actual ($f^{k,Actual}_{j,t}$), calculated based on the October WEO for the following year, or forecast ($f^{k,Forecast}_{j,t}$), calculated based on the October WEO for the current year. For instance, the *forecast* of fiscal spending for 2015 is taken from the fiscal variable in the October 2015 WEO, and the *actual* fiscal variable is taken from the fiscal variable in the October 2016 WEO, for the year 2015.[5]

The unanticipated fiscal variable shock is the difference between the actual and the forecast fiscal variable, where the latter is based on the information set as of the October WEO of the current year. This mitigates the anticipation effect, in which agents in the economy change their consumption and investment behavior based on the news about future fiscal policies for the rest of the year.

5. This formulation of a fiscal shock, FShock, is the difference in the level of the fiscal variable divided by the previous year's GDP. It is analogous to the shock used by Furceri and Li (2017) because:

$$\begin{aligned}
\text{FShock}^k_{j,t} &= \left(f^{k,Actual}_{j,t} - f^{k,Forecast}_{j,t}\right) = \frac{F^{k,Actual}_{j,t}}{Y_{j,t-1}} - \frac{F^{k,Forecast}_{j,t}}{Y_{j,t-1}} \\
&= \frac{F^{k,Actual}_{j,t} - F^k_{j,t-1}}{Y_{j,t-1}} - \frac{F^{k,Forecast}_{j,t} - F^k_{j,t-1}}{Y_{j,t-1}} \\
&= \frac{\Delta F^{k,Actual}_{j,t}}{Y_{j,t-1}} - \frac{\Delta F^{k,Forecast}_{j,t}}{Y_{j,t-1}}.
\end{aligned}$$

This is because whatever agents in the economy have anticipated, given the information set as of October, is already embedded in the forecast of fiscal variables.

By using the forecast of fiscal variables in the October WEO of the same year, we also lower the endogenous response of fiscal policy to the state of the economy in annual data. While the government could still change government consumption or investment in response to the state of the economy, our framework imposes the same assumption used by Blanchard and Perotti (2002), in that fiscal variables do not correspond contemporaneously to the state of the economy within the final quarter of the year (that is, between October and December). Since policymakers in many small states generally have access to fewer timely indicators to learn about the state of the economy than in larger economies, this timing assumption can be even more plausible in small states than in larger countries.

While our framework lowers endogeneity, it does not fully remove it. Government spending in many small states could respond to the state of the economy, for instance, by cutting spending in response to lower tax revenues arising from slow growth. However, as a check, we control for tax revenues and find that our results are robust.

Our forecast error approach employs the local projection method following Jordà (2005), in a similar spirit to Auerbach and Gorodnichenko (2013). The growth impacts of fiscal shocks are estimated using the following baseline specification:

$$y_{j,t+h,t-1} = \alpha_j^h + \gamma_y^h + \beta_I^h \mathrm{FShock}_{j,t}^I + \beta_C^h \mathrm{FShock}_{j,t}^C + \delta^h \mathbf{X}_{j,t} + \varepsilon_{I,t}^h, \tag{2}$$

where $y_{j,t+h,t-1}$ is the GDP growth rate between year $t-1$ and $t+h$ for country j; α_j is a country-specific fixed effect capturing factors that are time invariant; γ_t is the time fixed effect capturing global factors (for example, commodity price movements) that affect a country's growth in year t; $\mathrm{FShock}_{j,t}^I$ and $\mathrm{FShock}_{j,t}^C$ are the unanticipated fiscal variable shocks as a percentage of GDP for government investment and government consumption, respectively; and $\mathbf{X}_{i,t}$ is the set of control variables, including two lags of the GDP growth rate, two lags of each fiscal variable (in levels) as a percentage of GDP, the cumulative future fiscal variable shocks between years $t+1$ and $t+h$ (that is, $\sum_{l=1}^{h} \mathrm{FShock}_{j,t+l}^k$ for each type $k \in \{I, C\}$), and a natural disaster variable that captures the damages due to natural disasters as a percentage of GDP.[6]

6. In our baseline specification, the fiscal variables are divided by the previous year's GDP. However, the results are robust to a specification in which the fiscal variables are divided by the current year's GDP.

We include the cumulative future fiscal variable shocks occurring within the forecast horizon between t and $t + h$, $\sum_{l=1}^{h} \text{FShock}_{j,t+l}^{k}$, to avoid the biases pointed out by Teulings and Zubanov (2014).

Data

We use annual data for 1990–2017 from the IMF's *World Economic Outlook* (WEO) database. Since our local projection method uses two lags and four leads, our effective sample for estimation is from 1992 to 2013. Moreover, the panel is unbalanced due to the unavailability of data on fiscal variables for some periods for most countries. For our main empirical analysis, our small state sample is based on the IMF definition of small states (thirty-four countries). We further limit our sample by excluding some countries based on (1) insufficient data, (2) unreliable data (for example, negative government investment as a percentage of GDP), or (3) extremely large variance in government investment shocks, government consumption shocks, or GDP growth rates.[7] These restrictions reduce the number of small states for our empirical work to twenty-three, consisting of five countries from Africa, six from Asia, eleven from the Caribbean, and one from Europe (see table A1 in the online appendix).[8]

We use the October 2018 WEO to calculate the real GDP growth rate based on the real GDP series, *ngdp_r*. This is to avoid any possible measurement errors that may arise from data revision and updates of the compilation methodology. We then use the vintage IMF WEO database to calculate relevant variables. Government investment uses the series *gcek* prior to 2010 and

7. Based on the first two elimination criteria, East Timor, Maldives, Nauru, Palau, and Saint Lucia were excluded from the sample. For the third elimination criterion, non-Caribbean countries were dropped from the sample if (a) a standard deviation of government investment was above 15 percent of GDP or (b) a standard deviation of government consumption was higher than 20 percent of GDP. For Caribbean countries, observations were eliminated from the regression sample by putting outlier dummies if government investment shock was outside (–10, 10) percent of GDP or government consumption shock was outside (–15, 20) percent of GDP. These thresholds were calculated to include the ninety-eighth percentile of the respective variables. These outliers could reflect measurement errors and possible data revisions of government statistics or of the WEOs. Based on the third elimination criterion, Djibouti, Kiribati, Samoa, São Tomé and Príncipe, Tuvalu, and Vanuatu were eliminated. The paper's results are robust to large variations in these thresholds (not reported).

8. Supplementary material for this paper is available online at http://economia.lacea.org/contents.htm.

ggaan_t after 2010. Government consumption uses the current expenditure series *gcec* prior to 2000; thereafter it is total general government expense, *gge*, less interest payments, *ggei*.[9] Natural disaster damage data were obtained from EM-DAT.

Empirical Results

The empirical results show that government consumption has a small but positive impact on growth only in the short term, with almost no effect on growth over the medium term. Figure 1 (panel A) plots the baseline impacts of government consumption on GDP from equation 2. An increase in government consumption by 1.0 percent of GDP would increase output by about 0.3 percent on impact, which peaks in the following year at around 0.4. Over time, the impact of an increase in government consumption on the level of GDP decreases to zero. In other words, a dollar spent on government consumption will increase GDP by around 30 cents on impact and 40 cents in the second year, but it does not have a prolonged impact. Thus, government consumption has only a small and short-term impact on GDP.

On the other hand, government investment has a small effect on GDP at impact but a relatively large medium-term effect (figure 1, panel B). The effect of government investment on GDP rises gradually to around 0.2 percent of GDP in the second year and to around 0.9 percent in the fourth year. In other words, a dollar spent on investment increases GDP by 20 cents in the second year and by about 90 cents in the fourth year.[10]

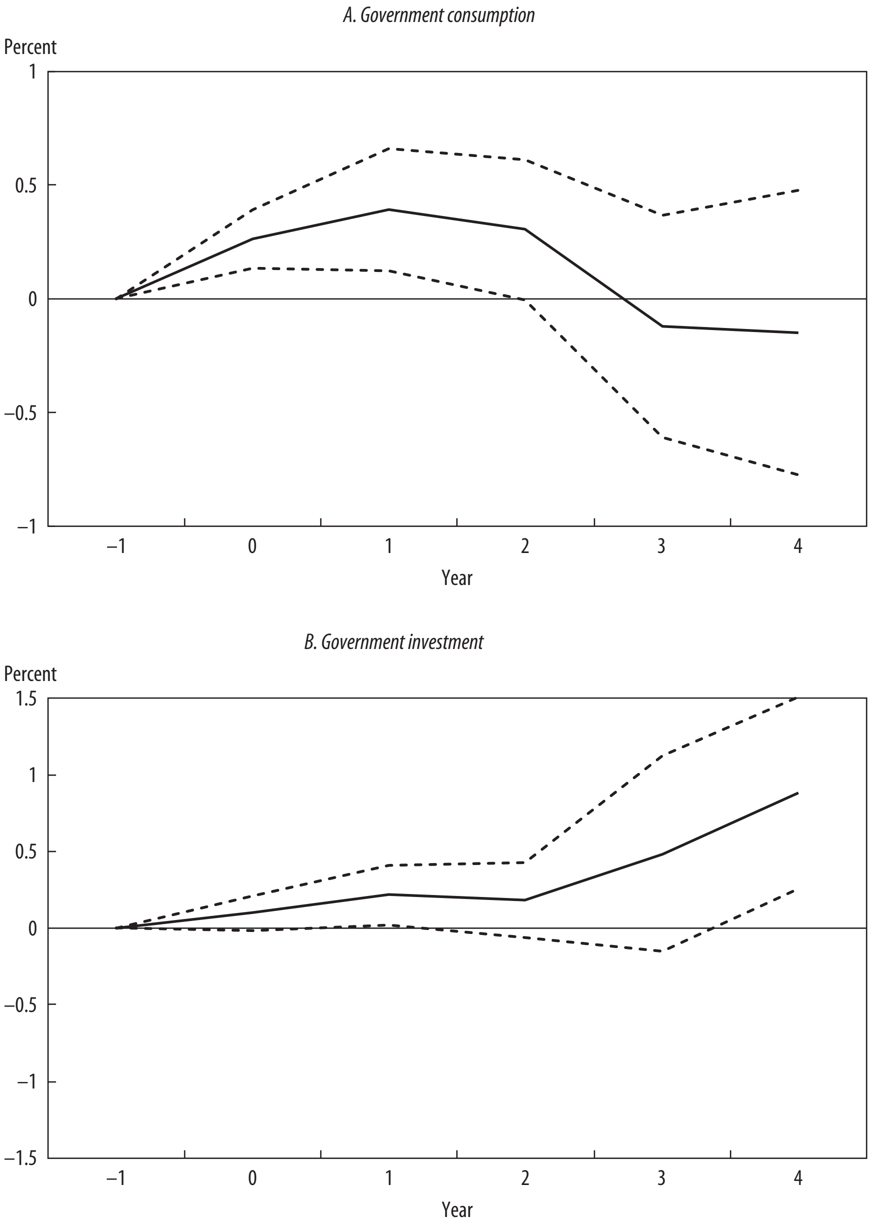
Expansion versus Consolidation

In this section, we investigate whether government spending has asymmetric effects on growth, depending on episodes of fiscal expansion or consolidation. In the local projection framework, this can be easily done by separating fiscal

9. As mentioned earlier, in the empirical part (only), government consumption is defined as government primary spending, which is government consumption plus transfers.

10. In our exercise, we do not calculate multipliers (as in Ramey and Zubairy, 2018) by dividing the cumulative changes in output by the cumulative change in the fiscal variable. We instead control future fiscal shocks in our regression and estimate the impact on GDP from the initial fiscal shock. Our fiscal multipliers are defined on the level of GDP in each period. We see this approach as more straightforward for calculating cumulative effects. We follow this definition throughout the paper, in both our empirical and GIMF models.

FIGURE 1. Empirical Results



Source: Authors' estimates.

Note: The figures show the response of GDP to a positive shock in government consumption and government investment equivalent to 1 percent of GDP, together with the 90 percent confidence intervals.

TABLE 1. Consolidation versus Expansion and Recession versus Boom

Fiscal multiplier	*Baseline* *(1)*	*Consolidation* *(2)*	*Expansion* *(3)*	*Recession* *(4)*	*Boom* *(5)*
Government consumption					
Impact	0.265***	0.392**	0.101	0.598***	0.059
	(0.079)	(0.212)	(0.190)	(0.257)	(0.180)
Peak	0.393***	0.842***	–0.139	0.793*	0.110
	(0.163)	(0.354)	(0.239)	(0.487)	(0.356)
Government investment					
Impact	0.0973	–0.0889	0.264***	0.814***	–0.414**
	(0.068)	(0.154)	(0.124)	(0.312)	(0.217)
Peak	0.882***	0.541	1.064***	1.537***	1.201**
	(0.380)	(0.770)	(0.421)	(0.701)	(0.652)

Source: Authors' estimates.

* $p < 0.125$; ** $p < 0.10$; *** $p < 0.05$.

Note: Standard errors clustered at the country level are in parentheses.

shocks into positive (expansionary) and negative (consolidation) episodes. We extend the specification in equation 1 as follows:

$$(3) \quad y_{j,t+h,t-1} = \alpha_j^h + \gamma_t^h + \beta_I^{h,Exp}\mathrm{FShock}_{j,t}^{I,Exp} + \beta_C^{h,Exp}\mathrm{FShock}_{j,t}^{C,Exp} + \beta_I^{h,Cons}\mathrm{FShock}_{j,t}^{I,Cons} + \beta_C^{h,Cons}\mathrm{FShock}_{j,t}^{C,Cons} + \delta^h \mathbf{X}_{j,t} + \varepsilon_{I,t}^h,$$

where $\mathrm{FShock}_{j,t}^{k,Exp}$ contains only positive (expansionary) fiscal shocks, defined as in equation 2, and $\mathrm{FShock}_{j,t}^{k,Cons}$ contains only negative (consolidation) fiscal shocks and is set to be zero otherwise.[11]

We find that the government's consumption multiplier is smaller for expansion episodes than for consolidation episodes (see table 1). This is consistent with the idea that an increase in government consumption, which would often result in an increase in public debt, may signal that fiscal tightening will happen in the near future, thus constraining the impact of fiscal expansion (Ilzetzki, Mendoza, and Végh, 2013). When a government increases its consumption, it does not boost GDP by much, either at impact or in the medium term.

11. As we divide the sample into two cases, the precision of the estimates becomes much weaker. To circumvent this problem, we also check the results for expansion versus consolidation and recession versus boom when we increase our sample size by extending our definition of small states to follow the World Bank's definition. This increases the sample from twenty-three to thirty-four countries. The results hold qualitatively true (see table A4 in the online appendix).

On the other hand, when the government reduces its consumption, it has a negative impact of around 0.4 percent on GDP at impact and 0.8 at the peak after one year.

Recession versus Boom

Similarly, we also investigate whether fiscal multipliers are larger in recessions than booms. We follow Auerbach and Gorodnichenko (2013) and modify equation 1 as follows:

$$(4)\quad y_{j,t+h,t-1} = \alpha_j^h + \gamma_t^h + \beta_I^{h,Recession} G(z_{it}) \text{FShock}_{j,t}^I + \beta_C^{h,Recession} G(z_{it}) \text{FShock}_{j,t}^C + \beta_I^{h,Boom} \left[1 - G(z_{it})\right] FShock_{j,t}^I + \beta_I^{h,Boom} \left[1 - G(z_{it})\right] \text{FShock}_{j,t}^C + \delta^h \mathbf{X}_{j,t} + \varepsilon_{I,t}^h,$$

where $G(z_{it}) = \exp(-\gamma z_{it})/[1 + \exp(-\gamma z_{it})]$, $\gamma > 0$, is a smooth transition function to give weights of the degree of recession for observations; and z_{it} is an indicator for the business cycle (in this case, the GDP growth rate) normalized to have zero mean and unit variance.[12]

Similar to previous studies (for example, Auerbach and Gorodnichenko, 2013), we find that both consumption and investment have a larger multiplier during recessions than booms (see table 1). For instance, while the government's consumption has a multiplier of 0.6 on impact during recessions, it does not have any notable effect on GDP during booms. In addition, while government investment has a large fiscal multiplier during recessions at around 0.8 on impact during recessions, it has a negative fiscal multiplier during booms.

Robustness Checks

We conducted a battery of robustness checks, including estimating equation 2 using just country fixed and time fixed effects; adding lagged variables, natural disasters, and future fiscal shocks; controlling for the terms of trade, net exports, government tax, and government revenue; and running separate regressions for government consumption and government investment (see table A3 in the

12. As in Aurbach and Gorodnichenko (2013), we set $\gamma = 1.5$. The results are robust to alternative values of γ.

online appendix). The results are robust to all these changes. The results are also robust to changes in control variables, such as a lag of the annual change in the actual fiscal variable (as in Auerbach and Gorodnichenko, 2013), lags of fiscal shocks, or combinations of these lagged fiscal variables. We also find that the results are robust to threshold values for classifying outliers, using trend GDP instead of actual GDP to divide variables, and using the previous year's WEO data (instead of the current year's) to obtain the fiscal variable forecast to construct fiscal variable shocks.

For robustness checks, we also conducted our analysis for small states based on the World Bank's definition (table A4 in the online appendix). After we perform the exclusion procedure described above, the sample includes thirty-four of the fifty small states under this definition, including ten countries in Africa, seven in Asia, twelve in the Caribbean, and five in Europe (see table A2 in the online appendix). Based on this sample, we find that the impact multiplier was 0.2 for government consumption, compared with the baseline multipliers of 0.3. For government investment, we find five-year multipliers of 0.6 for the larger sample, versus the baseline multipliers of 0.9. The five-year multipliers for government consumption and the impact multipliers for government investment are not statistically different from zero for either sample.

We also estimated multipliers for expansions and recessions for initially highly indebted countries (defined as having government debt of more than 70 percent of GDP), but the coefficients for these countries are similar to the baseline specification. This result should be viewed with caution, but it is not surprising, insofar as fourteen of the twenty-three countries in our sample are highly indebted. Our results using the DSGE model, discussed below, show that a higher level of government debt reduces fiscal multipliers.

Concerns for Monetary Policy

One may be concerned about potential omitted variables related to monetary policy (domestic and abroad). However, the majority of countries in our sample—all except Mauritius and Seychelles—do not have a floating exchange rate regime or an independent monetary policy (see table A1 for the exchange rate regimes in our sample). Furthermore, tests using the interest rate data for the limited subsample of eight countries for which data are available confirm that our fiscal shock measures are not correlated with interest rates in a statistically significant way. Finally, monetary policy abroad, such as U.S. monetary policy, is captured by our time fixed effects.

Comparison to Structural Vector Autoregression (SVAR)

Despite the caveats related to using structural vector autoregression (SVAR), which we discuss later, we performed regressions using this model and compared the results with our estimates based on the local projection method. Figure 2 shows the impulse responses to shocks in government consumption and investment from panel SVARs for different lags relative to our local projection method estimates. An SVAR with two-year lags mimics our local projection method in terms of the number of lags on the right-hand-side variables. An SVAR with four-year lags is shown to mimic four-year leads of our left-hand-side variable in the local projection method (namely, real GDP growth rates). Lag selection criterion prefers four-year lags over two-year lags for the SVAR estimates. We confirm that the SVAR results are generally within the 90 percent confidence interval of our local projection estimates and show similar patterns in responses to shocks in both government consumption and investment: the impact of government consumption shock is short-lived and dies out over the medium term, while the impact of a government investment shock has a larger impact in the long term.[13]

Summary of Empirical Results

The main results of our empirical portion are summarized in figure 3. Government consumption has a short-term impact multiplier of around 0.3 but a negligible medium-term impact on growth. In contrast, government investment has a small impact multiplier but a relatively large medium-term multiplier of around 0.9 on output.

DSGE Model: The Global Integrated Monetary and Fiscal Model (GIMF)

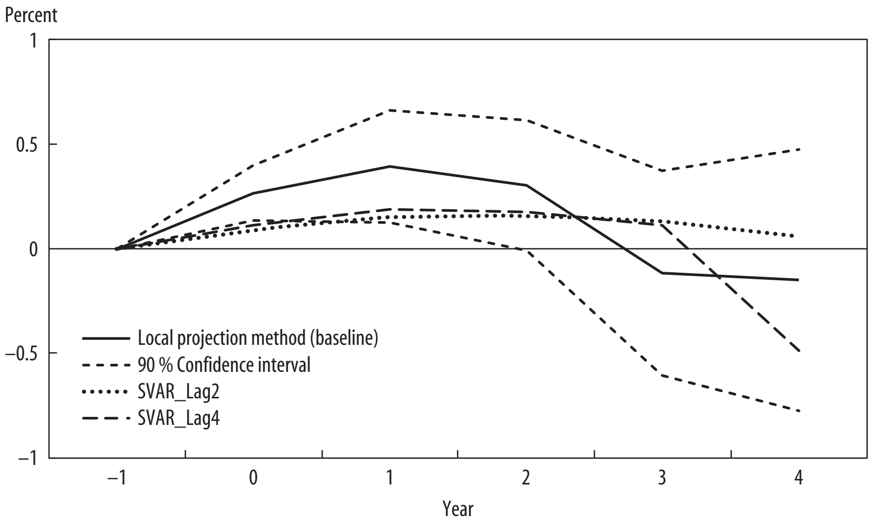
Our DSGE model is based on the IMF global integrated monetary and fiscal (GIMF) model.[14] This is an open economy model, in which Ricardian equivalence does not hold for various reasons. These include the model's feature

13. The fiscal multiplier estimates from SVAR and local projection methods are different even on impact ($h = 0$) for various reasons. The local projection method includes future fiscal shocks and government consumption and investment as a percentage of GDP instead of lagged fiscal shocks as in the SVAR. The local projection method also mechanically has a slightly shorter time dimension because its dependent variable is the real GDP growth rate up to the four-year horizon.

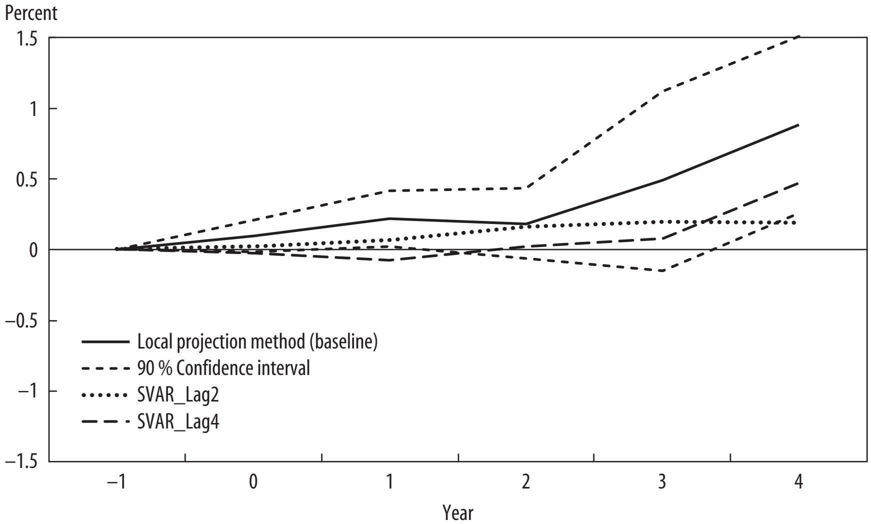
14. For detailed documentation on the structure of the model, see Kumhof and others (2010). Some explanations of the model are also borrowed from Leigh (2008).

FIGURE 2. SVAR Estimates

A. Government consumption



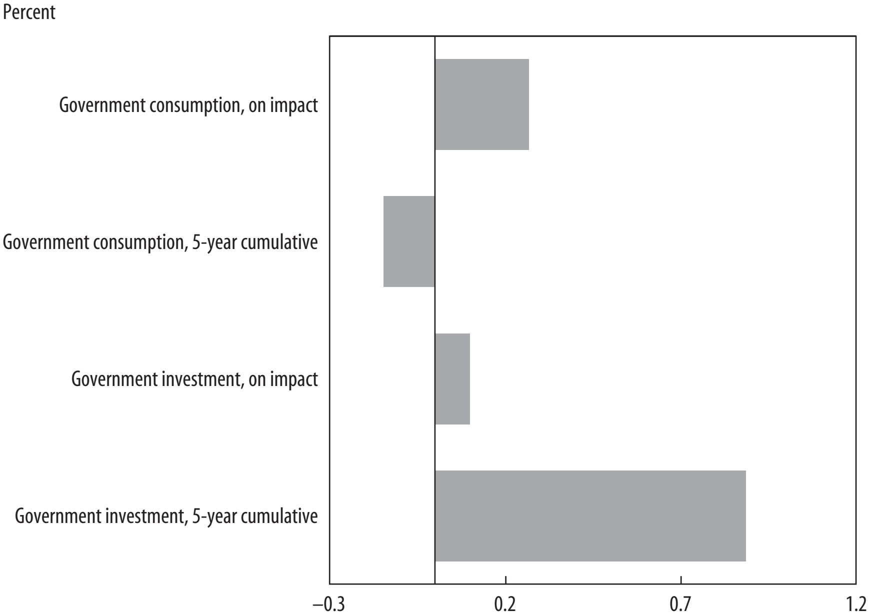
B. Government investment



Source: Authors' estimates.

Note: The figures show the response of GDP to a positive shock in government consumption and government investment equivalent to 1 percent of GDP, together with the 90 percent confidence intervals. SVAR_Lag2: panel SVAR with two lags. SVAR_Lag4: panel SVAR with four lags.

FIGURE 3. GDP Impact of Fiscal Expansion: Local Projection Method Results



Source: Authors' estimates.
Note: The figure shows the immediate and five-year cumulative effect on GDP of a fiscal expansion, equivalent to 1 percent of GDP, achieved through an increase in either consumption (government current primary spending) or investment.

of overlapping generations of agents with finite lifetimes, some of whom are also liquidity constrained. GIMF also has multiple real and nominal rigidities, including consumer habits that induce consumption persistence, investment adjustment costs that induce investment persistence, and import adjustment costs that induce spillover persistence from the policies of larger economies to the rest of the world.

GIMF relaxes the prevalent assumption in other DSGE models that all government spending is wasteful and does not contribute to aggregate supply. Instead, GIMF allows for productive public infrastructure spending that adds to the public capital stock and enhances the productivity of private factors of production.

The model's multiple non-Ricardian features, nominal and real rigidities, and fiscal and monetary policy reaction functions help produce plausible macroeconomic responses to changes in fiscal and monetary policy, as well as their spillover across economies. It is widely used to conduct policy analysis in IMF flagship publications.

Model

GIMF is a multicountry dynamic stochastic general equilibrium (DSGE) model with optimizing behavior by households and firms and full intertemporal stock-flow accounting. Frictions in the form of sticky prices and wages, real adjustment costs, and liquidity-constrained households, along with finite planning horizons for households, imply an important role in GIMF for monetary and fiscal policy in economic stabilization.

The assumption of finite horizons separates GIMF from standard monetary DSGE models and allows it to have well-defined steady states where countries can be long-run debtors or creditors. This allows users to study the transition from one steady state to another where fiscal policy and private saving behavior play a critical role in both the dynamics and long-run comparative statics.[15]

The non-Ricardian features of the model provide nonneutrality in both spending-based and revenue-based fiscal policy measures. In particular, fiscal policy can stimulate the level of economic activity in the short run, but sustained government deficits crowd out private investment and net foreign assets in the long run.[16] Sustained fiscal deficits in large economies can also lead to a higher world real interest rate, which is endogenous.

Asset markets are incomplete in the model. Government debt is only held domestically, in the form of nominal, noncontingent, one-period bonds denominated in domestic currency. The only assets traded internationally are nominal, noncontingent, one-period bonds denominated in U.S. dollars, which can be issued by the U.S. government and by private agents in any region. Firms are owned domestically. Equity is not traded in domestic financial markets; instead, households receive lump-sum dividend payments.

Firms employ capital and labor to produce tradable and nontradable intermediate goods. There is a financial sector à la Bernanke, Gertler, and Gilchrist (1999) that incorporates a procyclical financial accelerator, with the cost of external finance for firms rising with their indebtedness.

15. See Blanchard (1985) for the basic theoretical building blocks of such DSGE models and Kumhof and Laxton (2007, 2009a) for more detailed explanations of the fiscal policy implications of the GIMF model.

16. Coenen and others (2010) show that GIMF fiscal multipliers for temporary shocks are similar to standard monetary business cycle models, but GIMF can handle a much broader array of permanent shocks that can be used to study transitions from one steady state to another caused by permanent changes in the level of government debt.

GIMF is a multiregion model, encompassing the entire world economy, that explicitly models all the bilateral trade flows and their relative prices for each region, including exchange rates. The version used in this paper comprises three regions: a small state, the United States, and the rest of the world. The international linkages in the model allow the analysis of policy spillovers at the regional and global levels.

HOUSEHOLD SECTOR. There are two types of households, both of which consume goods and supply labor. First, there are households with overlapping generations (OLG) that optimize their borrowing and saving decisions over a twenty-year planning horizon. The first-order condition of their consumption-leisure choice sets their consumption relative to leisure proportional to the real disposable wage and the elasticity of labor supply. Second, there are liquidity-constrained households (LIQ), which do not save and have no access to credit. Both types of households pay direct taxes on labor income, indirect taxes on consumer spending, and a lump-sum tax.

Once we aggregate across households, we get the following condition:

$$\frac{c_t^{OLG}}{N\left(1-\psi\right)-l_t^{OLG}}=\frac{\eta^{OLG}}{\left(1-\eta^{OLG}\right)}w_t\frac{\left(1-\tau_{L,t}\right)}{\left(p_t^R-p_t^C\tau_{c,t}\right)},$$

where c_t^{OLG} is the per capita consumption of OLG households, η^{OLG} is the share of consumption versus labor in utility, N is the number of countries in the model, ψ is the share of liquidity-constrained households, l_t^{OLG} is the labor supply, w_t is the real wage, $\tau_{L,t}$ is labor income tax, p_t^R is the relative price of retail goods, p_t^C is the relative marginal cost of retailors, and $\tau_{c,t}$ is a consumption tax.

OLG households save by acquiring domestic government bonds, international U.S. dollar bonds, and fixed-term deposits. They maximize their utility subject to their budget constraint. Aggregate consumption for these households is a function of financial wealth and the present discounted value of after-tax wage and investment income. The consumption of LIQ households is equal to their current net income. Therefore, by construction, their marginal propensity to consume out of current income is unity.[17] A high proportion of LIQ households in the population would imply large fiscal multipliers from temporary changes to taxes and transfer payments.

17. The liquidity-constrained consumers could also be interpreted more generally as hand-to-mouth consumers, which in other models are assumed to consume all of their income.

For OLG households with finite planning horizons, a tax cut has a short-run positive effect on output. When the cuts are matched with a tax increase in the future, to leave government debt unchanged in the long run, the short-run impact remains positive, as the change will tilt the time profile of consumption toward the present. In effect, OLG households discount future tax liabilities at a higher rate than the market rate of interest. Thus an increase in government debt today represents an increase in their wealth, because a share of the resulting higher taxes in the future is payable beyond their planning horizon. If the increase in government debt is permanent (that is, tax rates are assumed to rise sufficiently in the long run to stabilize the debt-to-GDP ratio by financing the higher interest burden), this will crowd out real private capital by raising real interest rates.[18]

Increases in the interest rate have a negative effect on consumption, mainly through the impact on the value of wealth. The intertemporal substitution effect from interest rate changes is moderate and has been calibrated to be consistent with the empirical evidence. The intertemporal elasticity of substitution determines the magnitude of the long-run crowding-out effects of government debt since it pins down how much real interest rates have to rise to encourage households to provide the required savings.

PRODUCTION SECTOR. Firms, which produce tradable and nontradable intermediate goods, are managed following the preferences of their owners, who are the finitely lived households. Therefore, firms also have finite planning horizons. The main substantive implication of this assumption is the presence of a substantial equity premium driven by impatience.[19] Firms are subject to nominal rigidities in price setting, as well as real adjustment costs in labor hiring and investment. Investment adjustment costs, $\Gamma_{I,t}$, are as follows:

$$\Gamma_{I,t} = \frac{\phi_{I,t}}{2}\left(\frac{I_t}{I_{t-1}} - g \cdot n\right)^2,$$

where $\phi_{I,t}$ is the adjustment cost parameter, I_t is investment, g is the gross technological growth rate, and n is the population growth rate.

18. For a more detailed description of the fiscal implications in GIMF, see Kumhof and Laxton (2007, 2009a, 2009b).

19. This feature would disappear if equity was assumed to be traded in financial markets. We find the assumption of myopic firm behavior, and the resulting equity premium, to be more plausible.

The first-order condition from the firms' investment decision making is captured by Tobin's q:

$$q_t = P_t^I + \left(\Gamma_{I,t} + I_t \frac{\partial \Gamma_{I,t}}{\partial I_t}\right) P_t^I - E_t \left\{ \tilde{F}_t^{t+1} I_{t+1} \frac{\partial \Gamma_{I,t+1}}{\partial I_t} P_{t+1}^I \right\},$$

where P_t^I is the price of investment and the rest of the right-hand side is the net adjustment cost, with $\Gamma_{I,t} + I_t(\partial \Gamma_{I,t}/\partial I_t)$ representing the marginal adjustment cost of investment in t and $E_t\{\tilde{F}_t^{t+1} I_{t+1}(\partial \Gamma_{I,t+1}/\partial I_t) P_{t+1}^I\}$ representing the resulting adjustment cost savings in $t + 1$ from investment in t. In the steady state, Tobin's q is equal to the price of investment. In the dynamics, investment accelerates when Tobin's q is higher than the price of investment (sticky prices).

Firms operate in monopolistically competitive markets, and thus goods' prices contain a markup over marginal cost. Exports are priced to the local destination market and imports are subject to quantity adjustment costs. There are also price adjustment costs, which lead to sticky prices.

Firms use public infrastructure (which is the government capital stock) as an input, in combination with tradable and nontradable intermediate goods. Therefore, government capital adds to the productivity of the economy.

Firms also pay capital income taxes to governments and wages and dividends to households.

Retained earnings are insufficient to fully finance investment, so firms must borrow from financial intermediaries. If earnings fall below the minimum required to make the contracted interest payments, the financial intermediaries take over the firm's capital stock, less any auditing and bankruptcy costs, and redistribute it back to their depositors (households).

FINANCIAL SECTOR. GIMF contains a limited menu of financial assets. Government debt consists of one-period bonds denominated in domestic currency. Banks offer households one-period fixed-term deposits, which become their source of funds for loans to firms. These financial assets, as well as ownership of firms, are not tradable across borders. OLG households may, however, issue or purchase tradable U.S. dollar-denominated securities.

Banks pay a market rate of return on deposits and charge a risk premium on loans. Because of the costs of bankruptcy (capital can only be liquidated at a discount), the lending rate includes an external financing premium, which varies directly with the debt-to-equity (leverage) ratio—the financial accelerator effect. Nonlinearities imply steep increases in the risk premium for large negative shocks to net worth.

INTERNATIONAL DIMENSIONS AND SPILLOVERS. All bilateral trade flows are explicitly modeled, as are the relative prices for each region, including exchange rates. These flows include the export and import of intermediate and final goods. They are calibrated in the steady state to match the flows observed in the recent data. International linkages are driven by the global saving and investment decisions, a by-product of consumers' finite horizons. This leads to uniquely defined current account balances and net foreign asset positions for each region. Since asset markets are incomplete, net foreign asset positions are represented by nominal noncontingent one-period bonds denominated in U.S. dollars. A risk-adjusted uncovered interest rate parity sets the return on holding domestic bonds equal to holding the international bond, accounting for exchange rate risk and any other risk premiums.

Because of the importance of risk premiums in emerging markets and their possible relationship with fiscal policy, the model includes an endogenous country-specific risk premium. In particular, the risk premium on the interest paid on domestic debt, denoted by ρ, enters the risk-adjusted uncovered interest parity (UIP) equation for foreign currency bonds as follows:

$$i_t = i_t^{RW} E_t \varepsilon_{t+1} \left(1 + \rho_t\right),$$

where i_t^{RW} is the gross nominal interest rate in the rest of the world, and ε_{t+1} denotes future gross nominal exchange rate depreciation.

The domestic risk premium ρ_t is assumed to have the following nonlinear form:

$$\rho_t = \delta_1 + \frac{\delta_2}{\left[\left(\mathrm{DEBT}/\mathrm{GDP}\right)^{\max} - \left(\mathrm{DEBT}_t/\mathrm{GDP}_t\right)\right]^{\delta_3}}.$$

If $\delta_2 = 0$, then the risk premium always equals the exogenous level δ_1, regardless of the level of the debt-to-GDP ratio (DEBT/GDP). If $\delta_2 > 0$, a decline in government debt reduces the risk premium. As the debt-to-GDP ratio rises toward its maximum level, the risk premium rises at an increasing rate. The assumption of an increasing slope is broadly consistent with empirical studies that find a positive linear relationship between the logarithm of the risk premium and the debt ratio, such as Arora and Cerisola (2001). The parameter $\delta_3 > 0$ determines the curvature of the risk premium function.

Along with the adjusted uncovered interest parity and long-term movements in the world real interest rate, the magnitude of international trade linkages is the main determinant of spillover effects from shocks in one region to other world regions.

FISCAL AND MONETARY POLICY. Fiscal policy is conducted using a variety of expenditure and tax instruments. Government spending may take the form of consumption or investment expenditure or lump sum transfers, to all households or targeted toward LIQ households. Revenues accrue from taxes on labor and corporate income, consumption taxes, and lump sum taxes. The model also allows for tariffs on imported goods to be a potential source of public revenue. Government investment augments public infrastructure, which depreciates at a constant rate over time.

The government determines how the fiscal balance-to-GDP ratio responds to excess tax revenue using a simple fiscal policy rule:

$$\frac{\mathrm{FBAL}_t}{\mathrm{GDP}_t} = \phi^* + d\left(\frac{\tau_t - \tau_t^*}{\mathrm{GDP}_t}\right),$$

where FBAL/GDP is the fiscal balance-to-GDP ratio. If the response parameter $d = 0$, the fiscal balance is kept equal to ϕ^* at all times. For example, if $d = 0$ and the economy experiences an upswing, with actual tax revenue τ exceeding steady-state tax revenue τ^*, the fiscal balance remains unchanged, and the excess tax revenue is spent. Such a response corresponds to a balanced budget rule and is here defined as procyclical. A response of $d < 0$ would also qualify as procyclical. As the response parameter d increases in the positive range, a greater share of the excess tax revenue is saved. When $d = 1$, a 1 percent of GDP increase in excess tax revenue translates into a 1 percent increase in the fiscal balance, a response consistent with a structural balance rule. The rule can be implemented by adjusting taxes or spending. A response of $d > 1$ implies that a 1 percent of GDP increase in excess tax revenue induces an improvement in the fiscal balance of more than 1 percent of GDP; this is, for the purposes of this paper, defined as countercyclical.

The fiscal policy rule ensures long-run sustainability while allowing for short-run countercyclical policies. Changes in labor and capital income taxes or other taxes, transfers, or spending instruments provide instruments to put the rule into effect. First, the fiscal rule ensures that in the long run, the ratio of the government debt to GDP—and hence the deficit-to-GDP ratio—eventually converge to their target levels. This excludes the possibility of sovereign default, as well as the risk that out-of-control financing requirements of the government will override monetary policy. Second, the rule allows for countercyclical fiscal policy as it embodies automatic stabilizers.

When conducting monetary policy, the central bank in GIMF uses an inflation-forecast-based interest rate rule in the spirit of a Taylor rule. The central bank varies the gap between the actual policy rate and the long-run equilibrium rate to achieve a stable target rate of inflation over time. However, for this paper and the case of small states, where the nominal anchor is the exchange rate, there is no role for monetary policy, and this equation is not part of the model.

Calibration

The three-economy version of the GIMF used in the simulations has been calibrated to replicate key macroeconomic ratios such as external openness, tax collection and composition, fiscal spending patterns, and trade relationships among a hypothetical small state, the United States, and an aggregate of the rest of the small state's trading partners. The hypothetical small state is calibrated to broadly represent an average small state in terms of imports and government debt in percent of GDP. Its initial level for both imports and government debt is set at 61 percent of GDP, which is the 2017 average for the small states in the sample. Table 2 provides a summary of the calibration values for important parameters used in the baseline of this paper.

Each period corresponds to one year. The hypothetical small state is assumed to comprise 0.001 percent of world GDP and to have a steady-state annual real GDP growth rate of 1.5 percent and an inflation rate of 4 percent. The United States and the rest of the world are assumed to have a steady-state annual growth rate of 1.5 and an annual inflation rate of 2 percent. Population in all three regions is assumed to grow at 1 percent per year, and the real interest rate in the United States and the rest of the world is assumed to be 4 percent per year in the steady state. The structural parameters regarding household preferences and firm technology are set following Kumhof and Laxton (2007). In particular, the parameters that govern the degree of household myopia, a key non-Ricardian feature of the model, are calibrated as follows. Households in all three regions are assumed to have a planning horizon of fifteen years, a probability of death of 6.7 percent per year, and a decline in life-cycle worker productivity of 5 percent per year. Half of the small state's households are assumed to be liquidity constrained. This proportion is larger

TABLE 2. GIMF Baseline Calibration Values

Region and variable	*Calibrated value*
Small states	
Real GDP growth rate (percent; annual)	1.5
Inflation rate (percent; annual)	7.0
Real gross interest rate (percent; annual)	4.0
Population growth rate (percent; annual)	1.0
Share of liquidity-constrained agents (percent)	50.0
Imports (percent of GDP)	61.0
Fiscal ratios (percent of GDP)	
Government consumption to GDP	20.0
Public investment to GDP	4.7
Tax revenue to GDP	22.5
Consumption taxes	7.5
Capital taxes	4.0
Labor taxes	8.0
Lump sum taxes	3.0
Government debt	61.0
Labor shares (percent)	55.0
Labor shares, nontradables (percent)	60.0
World and the United States	
Investment share (percent)	17.2
Population share in the world: Small state (percent)[a]	0.0
Population share in the world: United States (percent)	23.0
Population share in the world: Rest of the world (percent)	77.0

Source: Authors' assumptions and estimates.

a. Population shares reflect the importance of trading partners from the perspective of small states and do not reflect the population shares of the region in the world per se.

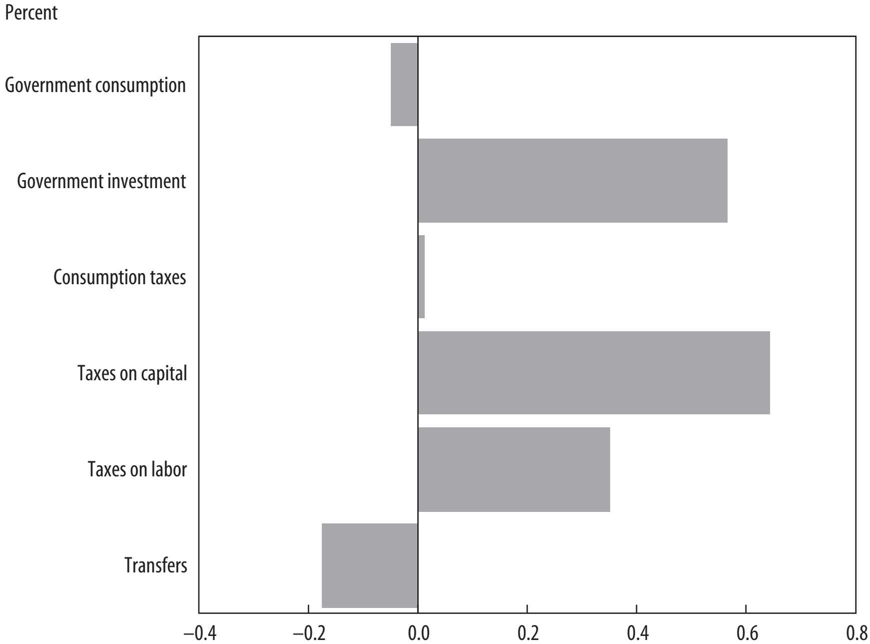
than the 33 percent that was assumed for the United States by Kumhof and Laxton (2007). Insofar as financial development is lower in small states than in the United States or other larger countries, a greater share of LIQ households in small states seems plausible.

The calibration of fiscal parameters, such as the ratios to GDP of government transfers, purchases of goods and services, and public investment, is broadly based on the averages of the small states.

GIMF Model Results

Our baseline multipliers are for public-debt-increasing shocks to fiscal policy variables that would increase the fiscal deficit permanently by 1 percent of GDP. The baseline assumes no monetary policy reaction to the fiscal shock

FIGURE 4. GDP Impact of Fiscal Expansion: GIMF Model Results



Source: Authors' estimates.

Note: The figure shows the five-year cumulative effect on GDP of a fiscal expansion, equivalent to 1 percent of GDP, achieved through an increase in consumption (government current primary spending), an increase in investment, a reduction in taxes, or an increase in transfers.

because most small states have either pegged exchange rates or an otherwise limited monetary policy.[20]

The five-year baseline fiscal multipliers are reported in figure 4. These are the effects of each shock on the level of GDP after five years. The government consumption multiplier is estimated at almost zero, meaning that after five years, the cumulative GDP effect of a fiscal expansion through increasing government consumption is almost zero. In other words, if the government of this small state expands its consumption such that its deficit is permanently higher by one percentage point of GDP, the economy would not enjoy any notable medium- or long-term effect of the policy on its GDP level. In contrast,

20. See tables A1 and A2 in the online appendix for the exchange rate classifications of our sample countries (from IMF, 2015).

TABLE 3. GDP Impact of Fiscal Expansion of 1 Percent of GDP: Time Profile
Percent; cumulative effect

	Year				
Variable	*0*	*1*	*2*	*3*	*4*
Government consumption	0.6	0.4	0.2	0.0	0.0
Government investment	0.7	0.6	0.5	0.5	0.6
Consumption taxes	0.4	0.3	0.2	0.1	0.0
Taxes on capital	0.5	0.6	0.5	0.5	0.6
Taxes on labor	0.3	0.4	0.4	0.4	0.4
Transfers	0.3	0.2	0.0	–0.1	–0.2

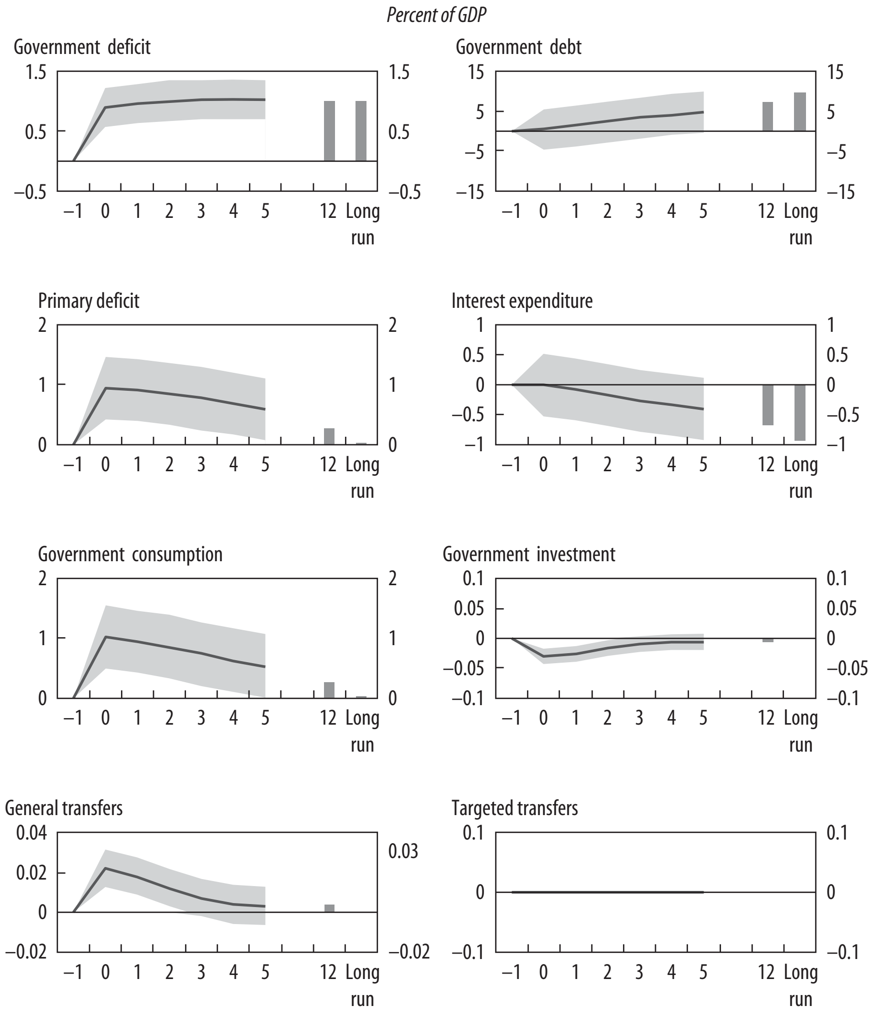
Source: Authors' estimates.

the five-year government investment multiplier is estimated at around 0.6. Thus, if the government increases its investment such that its deficit is permanently 1 percent of GDP higher, the economy will lose a cumulative 0.6 percent of its GDP over five years. Finally, five-year multipliers of expansion through reducing taxes range from about zero on consumption taxes to 0.4 on labor taxes and 0.6 on capital taxes.

Table 3 provides the path of multipliers from impact through five years. Multipliers are relatively larger at impact and decrease thereafter. In cases where fiscal expansion is achieved through the capital stock (that is, through government investment and taxes on capital), the multipliers increase over the medium term until they reach their steady-state levels. In the case of government consumption and consumption taxes, multipliers continue falling through the medium term and beyond until they reach zero. In the cases of labor taxes and transfers, the dynamics are much longer than the five-year horizon shown in table 3, but they also eventually reach zero (not shown).

To gain more insight into the baseline multipliers, we plot the dynamics of a set of important underlying fiscal and macroeconomic variables for a shock to government consumption (figures 5 and 6) and to government investment (figures 7 and 8). Figure 5 shows the government consumption shock and the resulting dynamics of fiscal ratios. As mentioned, the shock is calibrated to permanently increase the overall fiscal deficit by 1 percent of GDP, as graphed in panel A. The figure shows that government investment and transfers are virtually unchanged compared with the baseline. Insofar as the deficit is increased permanently, government debt rises on a declining trend compared with the baseline. This growth of government debt causes government interest expenditures to increase as well. Since the overall fiscal deficit is kept constant, the rising interest expenditures imply a deteriorating primary fiscal balance

FIGURE 5. Government Consumption: Fiscal Variables



Source: Authors' estimates.

Note: The figures show the response (deviation from steady state) of the different variables to a permanent expansion of government consumption equivalent to 1 percent of GDP. The long run is twenty years after the initial shock.

over time. This is a very important point for understanding the dynamics of the macroeconomic variables presented in figure 6, because the deteriorating primary fiscal balance acts similar to a stimulus for the economy.

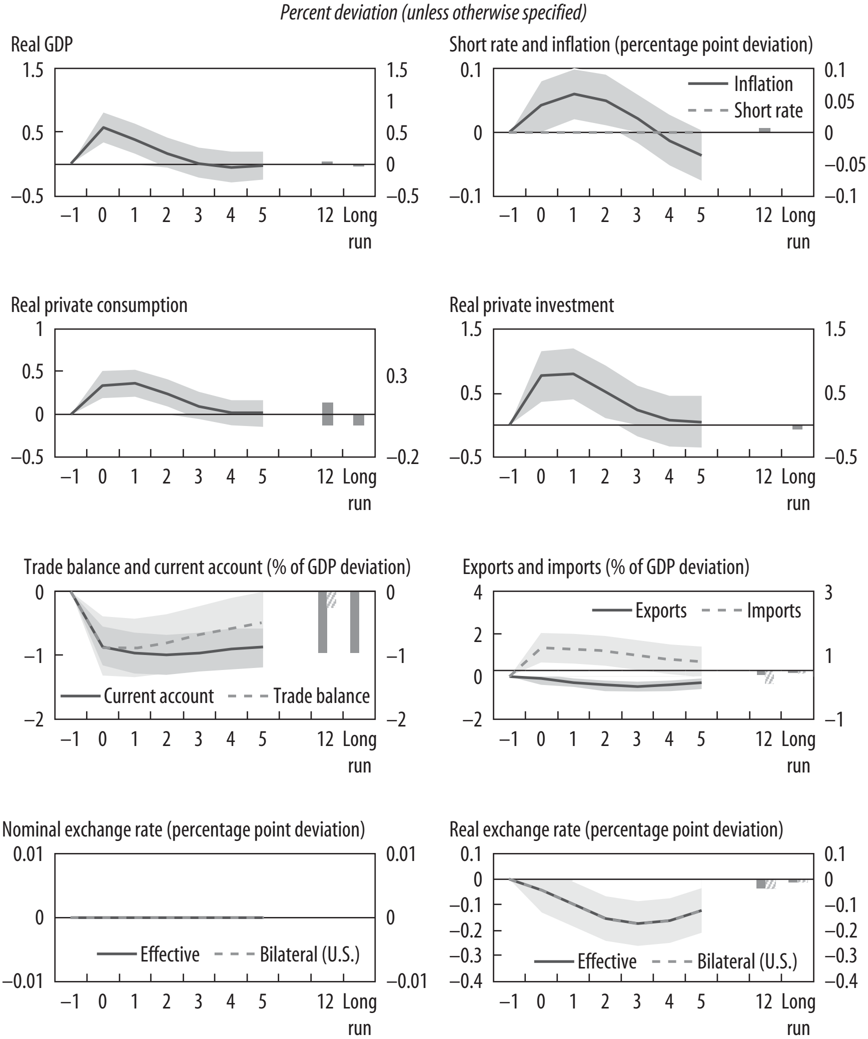
The first panel of figure 6 shows the evolution of real GDP. Because the shock was calibrated at a level to increase the deficit permanently by 1 percent of GDP, the resulting GDP path (relative to the steady state) can be interpreted as the fiscal multiplier path. This figure shows that the impact multiplier of a government consumption shock is about 0.6, but as time progresses, the multiplier shrinks, reaching zero after about four years. In the rest of this subsection, we describe the dynamics of various macroeconomic variables that are associated with this result, which are graphed in the remaining panels of the figure.

The positive government consumption shock causes private consumption and investment to increase at impact, as many consumers gain public jobs and many businesses obtain government contracts. However, as time goes by, private consumption and investment gradually return to their fundamental levels. This process is helped by the fact that the primary balance deteriorates after the impact. The expansion also leads to higher inflation and, with the nominal exchange rate broadly unchanged, results in a real exchange rate appreciation. This dampens exports somewhat and boosts imports. A larger boosting effect on imports is realized at impact because both government and private domestic demand expand. Over time, however, private demand deteriorates, so imports also partially decline.

Table 4 presents the contributions of different variables to growth. The first row shows the total impact on GDP (or the fiscal multiplier) over six years when the government expands its overall fiscal deficit by 1 percent of GDP through increased government consumption. Private consumption and investment also increase as a response to a positive government consumption shock. If there were no trade leakage, GDP would expand by around 1.3 percent. However, imports would also increase as a result of higher government and private demand. This trade leakage dampens the original impact of an increase in government consumption and brings down the overall GDP impact to around 0.6 percent. Over time, both consumption and investment decline, and the trade balance improves through an increase in exports and a decline in imports.

Figure 7 shows the government investment shock and the resulting dynamics of different fiscal variables. The shock is calibrated to permanently increase the overall fiscal deficit by 1 percent of GDP through government investment. Government investment is expanded, while government consumption

FIGURE 6. Government Consumption: Nonfiscal Macroeconomic Variables



Source: Authors' estimates.

Note: The figures show the response (deviation from steady state) of the different variables to a permanent expansion of government consumption equivalent to 1 percent of GDP. The long run is twenty years after the initial shock.

TABLE 4. Permanent Expansion of the Deficit by 1 percent of GDP using Higher Government Consumption

Percentage point contribution to GDP

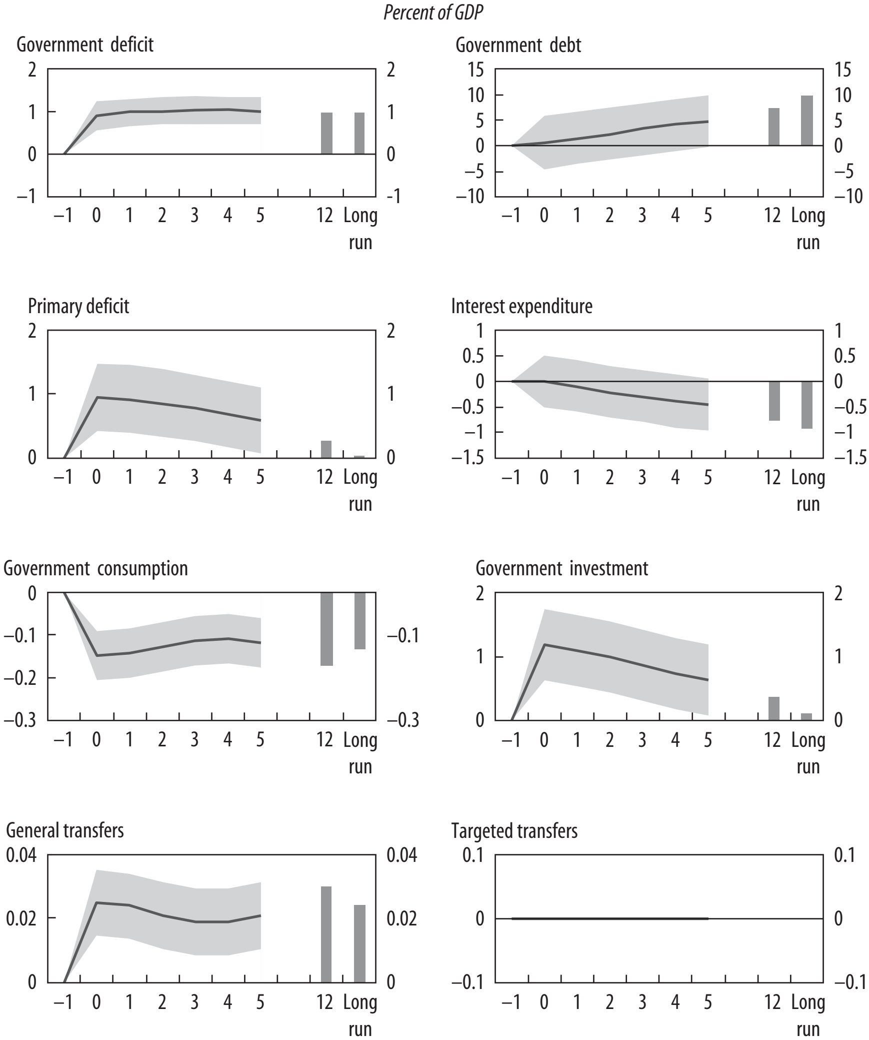
	Year				
Variable	*0*	*1*	*2*	*3*	*4*
Total impact on GDP	**0.6**	**0.4**	**0.2**	**0.0**	**0.0**
Private consumption	0.2	0.2	0.1	0.1	0.0
Private investment	0.1	0.1	0.1	0.0	0.0
Government spending	1.1	1.0	0.9	0.8	0.7
Government consumption	1.2	1.1	0.9	0.8	0.7
Government investment	0.0	0.0	0.0	0.0	0.0
Net exports	–0.9	–1.0	–1.0	–0.9	–0.7
Exports	–0.1	–0.2	–0.3	–0.3	–0.3
Imports	0.8	0.8	0.7	0.6	0.5

Source: Authors' estimates.

and transfers remain virtually unchanged through the steady state. Similar to the previous case of a decline in the overall fiscal balance through government consumption, government debt increases over time owing to a permanently higher fiscal deficit. Since the overall fiscal deficit is kept constant, the primary fiscal balance deteriorates over time with higher government debt.

Figure 8 plots the dynamics of macroeconomic variables in response to the permanent increase in the overall fiscal deficit by 1 percent of GDP. Given that the shock was calibrated at a level to increase the deficit permanently by 1 percent of GDP, the resulting GDP path (relative to the steady state) can be interpreted as the fiscal multiplier path. Similar to the case of government consumption, a 1 percent of GDP increase in the overall fiscal deficit positively affects private consumption and investment at impact, as many consumers gain public jobs and many businesses gain government contracts. Over time, both private consumption and investment decline, but they end up at higher steady-state levels because the positive effect of higher government investment is permanent. The expansion also increases inflation and appreciates the real exchange rate, which deteriorates the trade balance by dampening exports while boosting imports. Over time, private demand declines, and imports also partly decline. Unlike the case of government consumption, however, the increase in government investment also positively affects the capital stock in the economy and leads to higher production. Thus the increase in government investment has a more lasting impact on output.

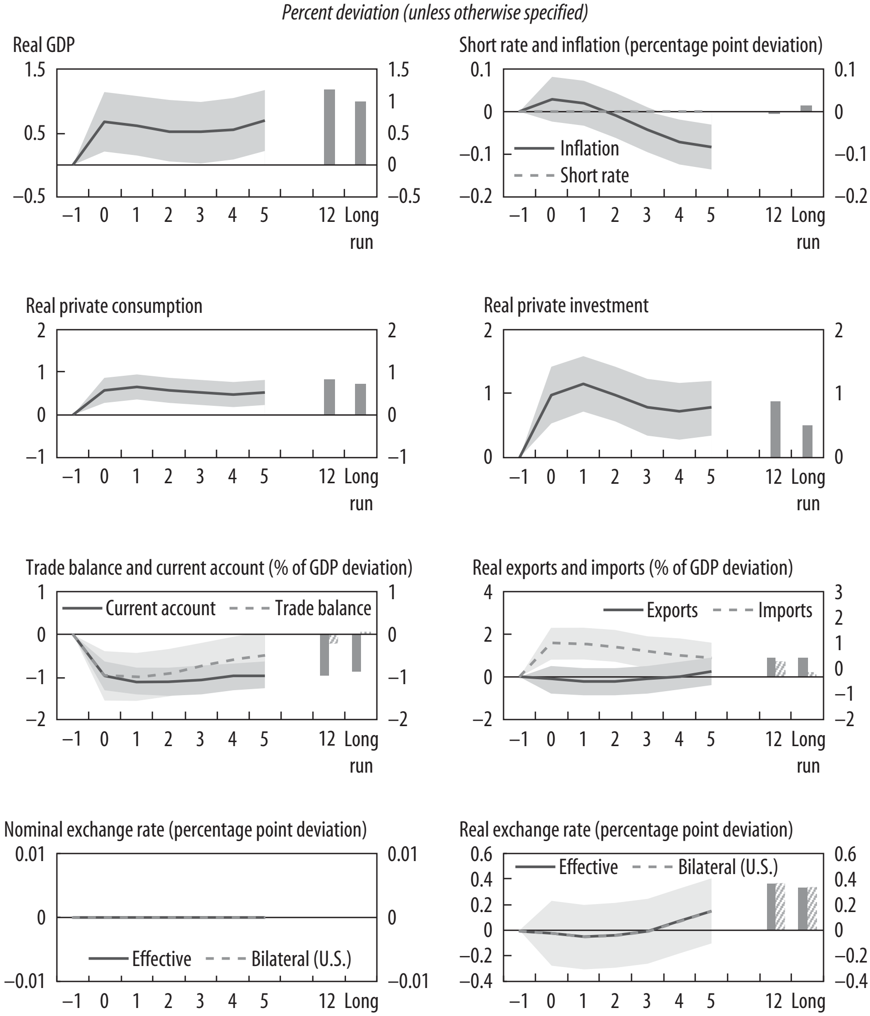
FIGURE 7. Government Investment: Fiscal Variables



Source: Authors' estimates.

Note: The figures show the response (deviation from steady state) of the different variables to a permanent expansion of government consumption equivalent to 1 percent of GDP. The long run is twenty years after the initial shock.

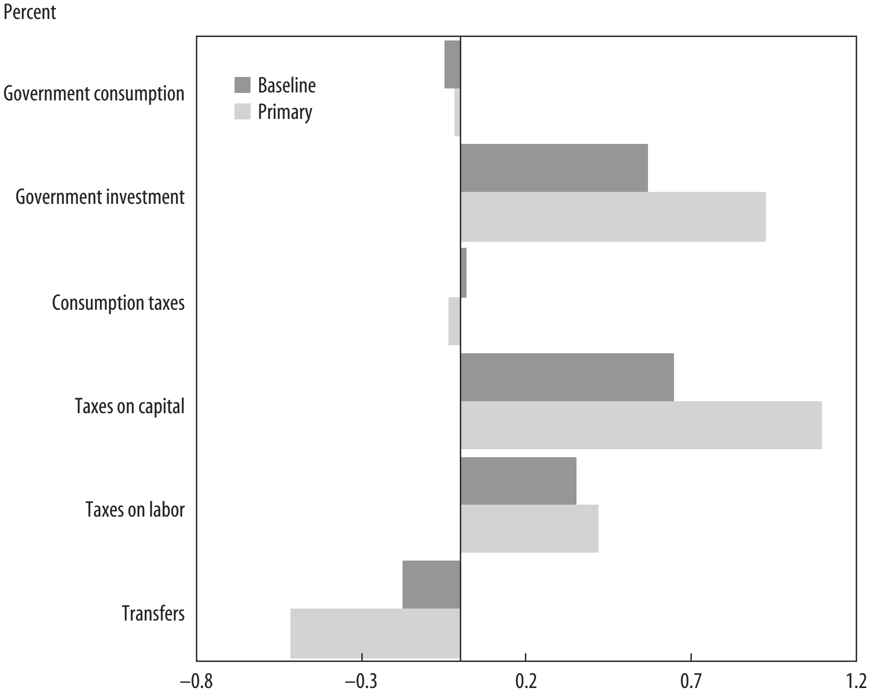
FIGURE 8. Government Investment: Nonfiscal Macroeconomic Variables



Source: Authors' estimates.

Note: The figures show the response (deviation from steady state) of the different variables to a permanent expansion of government consumption equivalent to 1 percent of GDP. The long run is twenty years after the initial shock.

FIGURE 9. Primary Deficit versus Baseline



Source: Authors' estimates.

Note: The figure shows the five-year cumulative effect on GDP of an increase in the overall deficit (baseline) versus an increase in the primary deficit (overall deficit minus interest), both equivalent to 1 percent of GDP, achieved through an increase in consumption (government current primary spending), an increase in investment, a reduction in taxes, or an increase in transfers.

In the very long term (well beyond our definition of twenty years for the long term), the output effect will return to zero as private investment is replaced by the increased public investment due to the permanent expansion.

Primary Balance Multipliers

In the baseline, the size of the policy shocks is always set such that they increase the overall deficit by 1 percent of GDP. In some cases, however, policymakers are interested in multipliers for a change in the primary deficit (that is, the overall deficit minus interest) by 1 percent of GDP. Figure 9 shows our models' results for these multipliers and compares them with the baseline multipliers.

As the figure shows, the primary balance multipliers are larger than the baseline multipliers. This is an intuitive result. With the expansion, government debt is on an upward path, which causes interest expenditures to increase over time. Baseline multipliers assume a constant overall deficit after the impact. Therefore, in the baseline, the primary deficit decreases over time to compensate for higher interest expenditures. In the case of primary multipliers, the primary deficit remains unchanged over time, at 1 percent of GDP higher than the steady state. This results in higher primary balance multipliers than the baseline.

Temporary Shock Multipliers

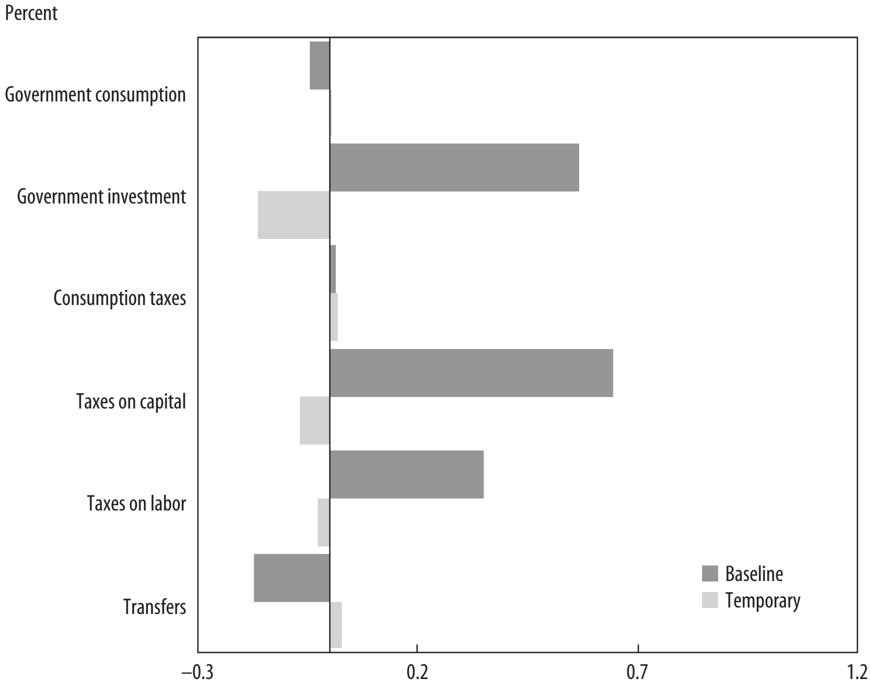
The baseline multipliers were estimated for a permanent expansion shock. In this subsection, we present the multipliers for a temporary expansion shock. In this exercise, fiscal policy variables are changed to increase the overall deficit in the first year by 1 percent of GDP, but the deficit returns back to the steady-state level in the following year. The overall deficit in all future years is kept unchanged relative to the steady state. Figure 10 shows the results and compares them with the baseline. Temporary multipliers are notably smaller than baseline multipliers. This is as expected because the present value of a temporary fiscal shock is much smaller than that of a permanent one with the same annual size. Also, some temporary multipliers are estimated with the wrong signs. This is because of various dynamics across variables in the model, and it is of little importance because of the very small size of the multipliers.

Multipliers following Natural Disasters

While our baseline GIMF estimation of fiscal multipliers in the previous sections has assumed that the small state starts at the steady state, in reality, many small states are often hit by natural disasters (such as hurricanes) that take them well out of their steady state. Following a natural disaster, fiscal policy is usually considered an important tool to bring the economy back toward its steady state. This section estimates fiscal multipliers after a natural disaster. We consider a natural disaster that destroys 10 percent of the country's GDP in the initial period, following which fiscal policy is implemented.

Figure 11 plots five-year cumulative GDP impacts of government consumption and investment in this post-natural-disaster economy. The fiscal stimulus from government consumption following a natural disaster is estimated to have a medium-term multiplier of close to 0.4. This is notably larger than in the baseline, which has a multiplier of almost zero. The medium-term government

FIGURE 10. Temporary Shock versus Baseline



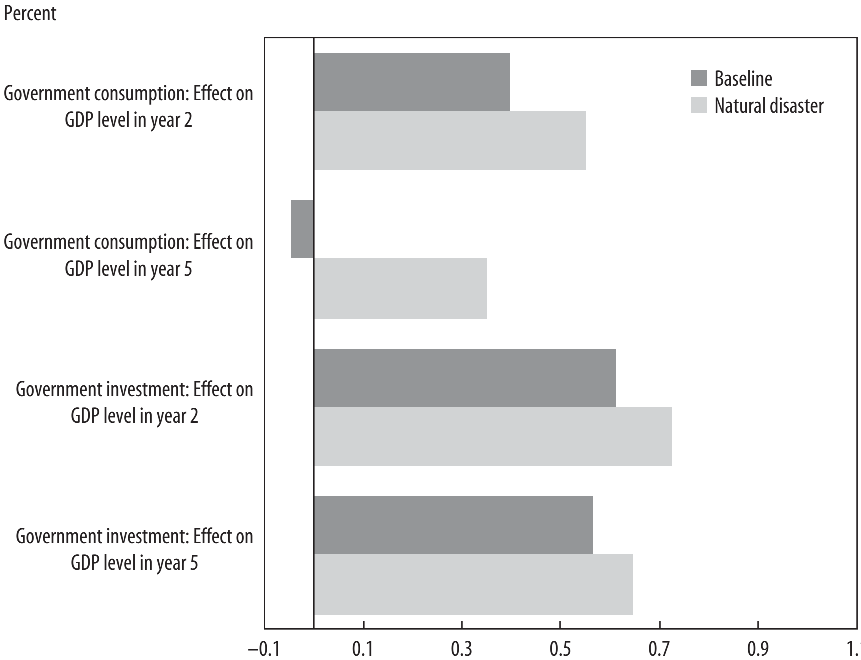
Source: Authors' estimates.

Note: The figure shows the five-year cumulative effect on GDP of a temporary expansion shock, in which the deficit returns back to the steady-state level in the following year, versus the baseline, both equivalent to 1 percent of GDP, achieved through an increase in consumption (government current primary spending), an increase in investment, a reduction in taxes, or an increase in transfers.

investment multiplier after a natural disaster is estimated at 0.7, slightly larger than the baseline. These results are intuitive because one expects to have larger multipliers when there is slack in the economy. The results are also consistent with our empirical results presented earlier, which found larger multipliers in recessions compared to booms.[21]

21. Our conclusions from the analysis of natural disasters may not be directly applicable to the COVID-19 pandemic owing to its differences from natural disasters. A natural disaster generally destroys physical capital but may not have a longlasting impact on consumer behavior. In contrast, a health disaster like the COVID-19 pandemic does not destroy existing physical capital in the economy, but it could have a longlasting impact on consumer behavior, such as a long-term decline in the demand for tourism. At the same time, the COVID-19 shock also entails important disruptions to the supply side (via lockdowns), which would suggest a more muted impact for fiscal policy at least in the containment phase.

FIGURE 11. Post-Natural Disaster versus Baseline



Source: Authors' estimates.

Note: The figure shows the five-year cumulative effect on GDP of a fiscal expansion following a natural disaster that destroys 10 percent of the country's GDP in the initial period, versus the baseline, both equivalent to 1 percent of GDP, achieved through an increase in consumption (government current primary spending), an increase in investment, a reduction in taxes, or an increase in transfers.

Sensitivity Analysis

The baseline and other previous sections were calibrated for a hypothetical small state with specific characteristics (table 2). Most notably, imports and government debt level of the baseline's small state were set at the average levels of all small states. Given the diversity of small states, in this section we provide a sensitivity analysis of the results with respect to three important country characteristics: the import share, the government debt level, and the share of LIQ households.

Figure 12 plots the GDP cost of fiscal consolidation in response to a 1 percent of GDP expansionary fiscal shock on government consumption, government investment, consumption tax, capital tax, labor tax, or transfers, both on impact and over a five-year horizon, for each of the three characteristics

FIGURE 12. Sensitivity Analysis

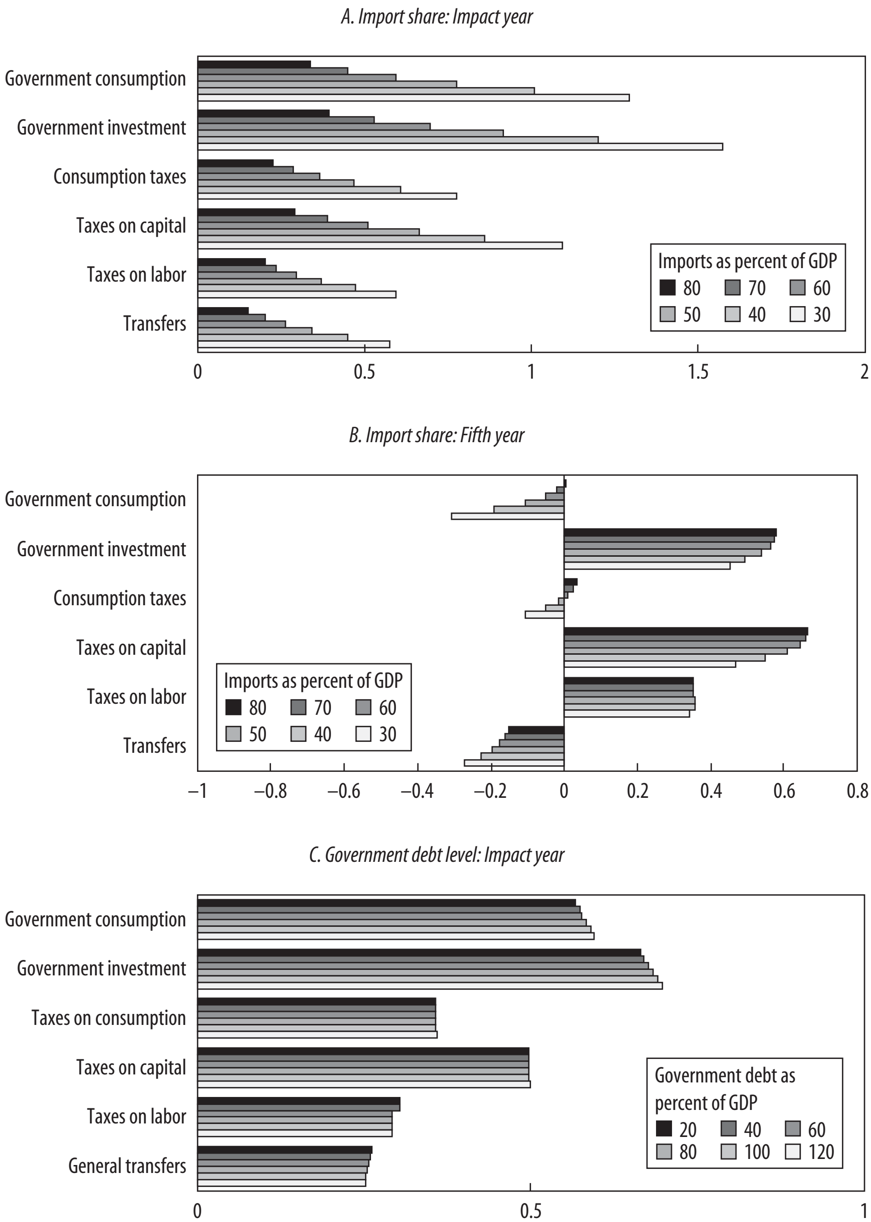
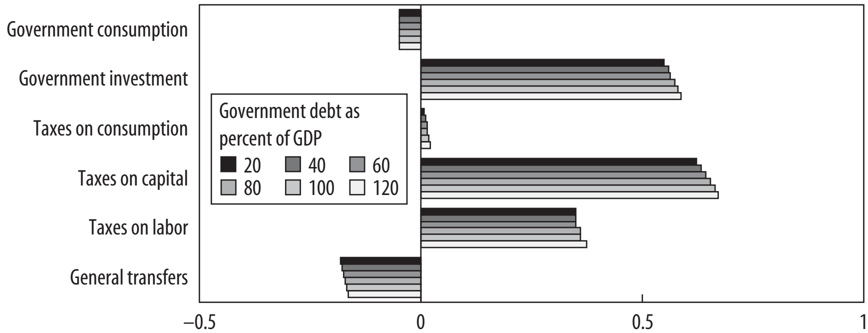
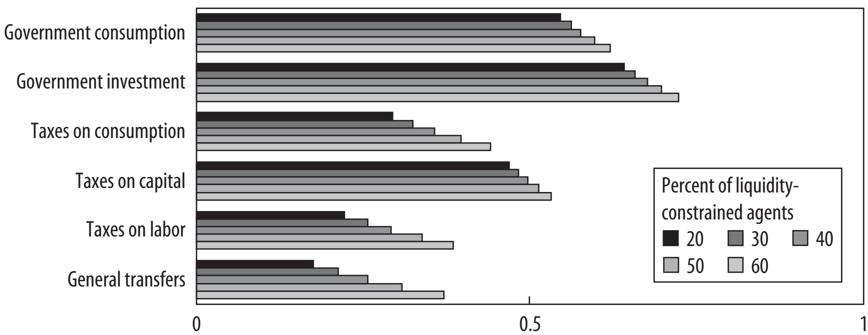


FIGURE 12. Sensitivity Analysis (*Continued*)

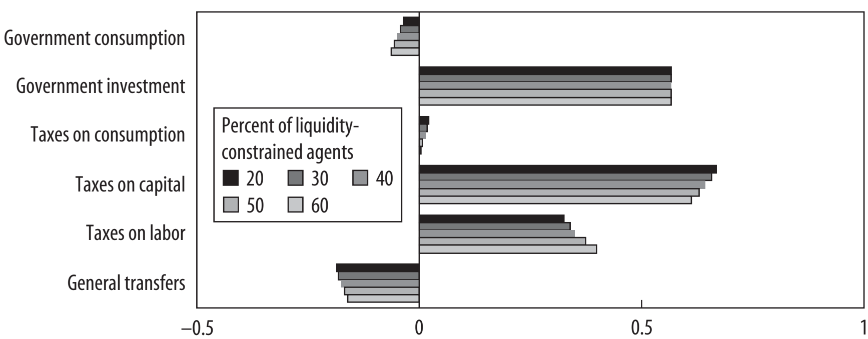
D. Government debt level: Fifth year



E. Share of liquidity-constrained households: Impact year



F. Share of liquidity-constrained households: Fifth year



Source: Authors' estimates.

Note: The figures show the one- and five-year cumulative effect on GDP of a fiscal expansion, equivalent to 1 percent of GDP, achieved through an increase in consumption (government current primary spending), an increase in investment, a reduction in taxes, or an increase in transfers, for different initial levels of the import share, government debt, and the share of liquidity-constrained households.

identified above. Panels A and B show a sensitivity analysis in which the import share is progressively raised from 30 to 80 percent of GDP. The higher the import share, the lower the fiscal multipliers, because the trade leakage is greater when import shares are higher. This holds for both on impact (one-year effect) and in the medium term (five-year effect).

Panels C and D of figure 12 plot the fiscal multipliers for different levels of government debt, ranging from 20 to 120 percent of GDP. The higher the government debt level, the higher the fiscal multipliers. This is because consolidation lowers the risk premium more for countries with higher debt levels and is thus more beneficial to those countries.

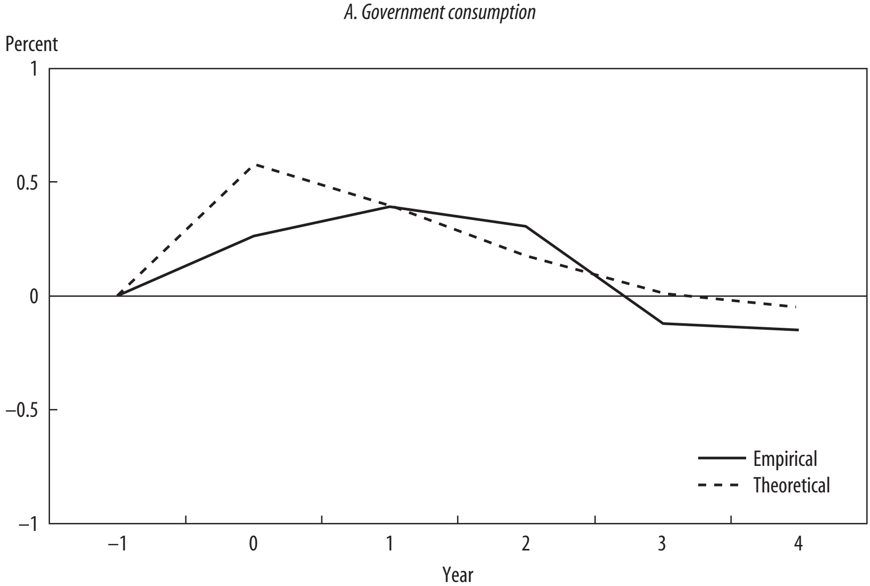
Lastly, panels E and F provide a sensitivity analysis for the share of LIQ households, with values between 20 and 60 percent of the population. Here again, the fiscal multipliers increase in step with the share of LIQ households. This reflects the fact that LIQ households have a hand-to-mouth consumption behavior and thus have a higher marginal propensity to consume, resulting in a larger fiscal multiplier.

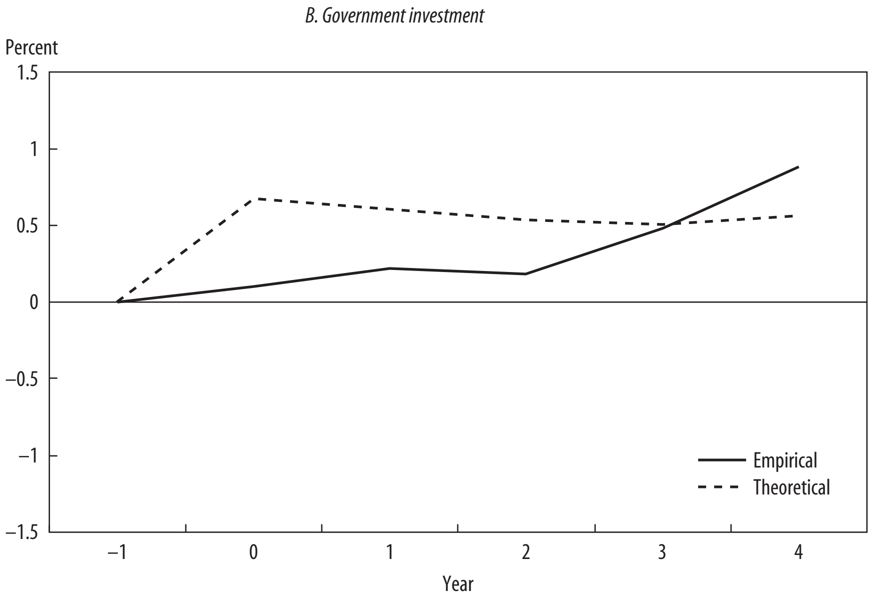
Comparing Empirical and Theoretical Multipliers

In this subsection, we compare the impulse responses to government consumption and investment shocks from our empirical and theoretical (GIMF) models. Figure 13 plots the impulse responses of GDP to a 1 percent of GDP shock in government consumption and investment across the two models. The dynamics of government consumption are very similar between the empirical and GIMF models. In case of government investment, however, our GIMF model results suggest that the shock has a sizable immediate impact, whereas our empirical results indicate that it does not have an immediate impact.

The differences in the initial impact of government investment between the theoretical model and the empirical results are partly attributable to a potentially higher than average import share in government investment. As the sensitivity analysis in figure 12 shows, the higher the import share, the smaller the fiscal multiplier on impact. This is because an increase in GDP resulting from an increase in government spending is partly offset by a decline in net exports (that is, an increase in imports). In reality, government investment could have a higher import share than government consumption. That is, small states must import a large share of capital goods for government investment projects from abroad, whereas a relatively larger share of government consumption goods can be produced domestically. This disproportionately higher import share of government investment, in turn, results in a smaller

FIGURE 13. Empirical versus Theoretical Results





Source: Authors' estimates.

Note: The figures show the response (deviation from steady state) of GDP to a permanent expansion of government consumption and government investment equivalent to 1 percent of GDP under our empirical (local projection) and theoretical (GIMF) models.

fiscal multiplier on impact, as suggested by our empirical results. Additionally, over the medium term, government investment increases the government capital stock both in the model and in the data, which results in higher output, thus showing a similar high medium-term impact of government investment both in the model and in the empirical results.

Comparisons with Previous Studies

This section reviews the existing relevant studies and compares their results with ours. There are only a few existing contributions in the literature that estimate fiscal multipliers for small states. We categorize them based on their methodologies, namely, SVAR, narrative approach, and DSGE model. In this section, we briefly explain these methodologies and provide reasons why our forecast error methodology is more plausible for estimating fiscal multipliers for small states.

When an SVAR is used to identify government spending shocks, as in Blanchard and Perotti (2002), it is assumed that government spending does not respond to a change in GDP within the contemporaneous period (for example, quarter or year). Gonzalez-Garcia, Lemus, and Mrkaic (2013) estimate an SVAR with panel quarterly data interpolated from annual data for Eastern Caribbean Currency Union (ECCU) countries. They find that the fiscal multiplier for government consumption is not statistically significantly different from zero, while the fiscal multiplier for government investment is slightly less than 0.4 after one year. Using interpolated quarterly data for fourteen Caribbean countries between 1990 and 2011, Narita (2014) estimates an SVAR and finds that impact multipliers for government consumption are 0.1–0.2 on impact and 0.0–0.3 in the medium term. Guy and Belgrave (2012) employ an SVAR approach to estimate fiscal multipliers for government expenditure for four Caribbean countries by interpolating annual data into quarterly data between 1980 and 2008.[22] They find that the fiscal multipliers for government expenditures are very small after one year, at 0.1, and range from a small negative to 0.3 over a six-year period. Neither Guy and Palgrave (2012) nor Narita (2014) distinguish government consumption from government investment.

The second approach, known as the narrative approach, uses the news and budget documents to identify unexpected fiscal spending shocks by dropping

22. The four Caribbean countries in their study are Barbados, Guyana, Jamaica, and Trinidad and Tobago.

the incidences of government spending increases in response to current or prospective economic conditions (for example, David and Leigh, 2018; Romer and Romer, 2010). Data are not available for most small states to conduct this approach.

Finally, Dodzin and Bai (2016) calibrate a DSGE model for Palau and Kiribati and estimate an impact government consumption multiplier of around 0.5. To the best of our knowledge, this is the only existing study that uses a DSGE model to estimate multipliers for small states.

Key Advantages of Our Empirical Approach Compared to the Existing Empirical Literature

Our approach has a number of advantages over the previous literature on small states. First, the sample in our study is much larger than the sample used in previous studies for small states, with twenty-three countries based on the IMF definition of small states and thirty-four countries based on the World Bank's definition. In contrast, the previous studies cited above used four small states (Guy and Belgrave, 2012), eight ECCU countries (Gonzalez-Garcia, Lemus, and Mrkaic, 2013), and fourteen Caribbean countries (Narita, 2014).

Second, unlike the SVAR approach used in the majority of the previous studies, which imposes a recursive structure on responses to shocks, the local projection method in this paper allows nonlinear responses of GDP to changes in government spending. In essence, on impact (when horizon h is equal to 0), the effect is the same under an SVAR approach and a local projection method because at horizon 0, the GDP equation from a local projection method is a restricted version of the GDP equation from a recursive SVAR, where the lagged fiscal policy has no effect on contemporaneous GDP. The difference arises after the impact, where a local projection method allows a nonlinear response. Moreover, we augment the simple local projection method to avoid bias by including future fiscal shocks, as pointed out by Teulings and Zubanov (2014).

Third, our forecast error approach helps avoid potential measurement errors arising from interpolating annual data, which is the only data frequency available for many small states, into a quarterly frequency. Studies that use an SVAR interpolate annual data to obtain quarterly data (for example, Narita, 2014). Such an approach relies on how good the interpolation is. Given that many small states do not have official quarterly GDP statistics, the interpolation method could generate severe measurement errors. Under the forecast error method based on the October WEO of the same year, our identification assumption for estimating fiscal multipliers is similar to the assumption used for SVARs run for quarterly data, in which fiscal variables do not respond to the state of the economy within a quarter (October to December). As a result, we also mitigate endogeneity issues, which could have been severe if we had simply estimated fiscal multipliers at an annual frequency without using a forecast error approach.

Fourth, the forecast error method can dampen the anticipation effect of the fiscal variable. While this methodology relies on a similar timing assumption as an SVAR estimation based on quarterly data (for example, Blanchard and Perotti, 2002), the forecast error approach mitigates the anticipation problem in which agents respond by changing their consumption and investment behavior before the actual realization of changes in government spending. Previous studies that estimate fiscal multipliers for small states using an SVAR model do not account for the foresight problem.[23]

Fifth, this is the first paper to estimate state-dependent fiscal multipliers for small states (namely, expansion versus consolidation and boom versus recession). None of the aforementioned papers do so (Gonzalez-Garcia, Lemus, and Mrkaic, 2013; Guy and Palgrave, 2012; Narita, 2014). Unlike SVARs, the local projection method allows the estimation of state-dependent fiscal multipliers. Our empirical results suggest that fiscal multipliers are much larger during recessions and consolidations than during booms and expansions. This highlights the importance of considering state dependency (as shown in table 1).

Last, this paper estimates government consumption and investment separately. With the exception of Gonzalez-Garcia, Lemus, and Mrkaic (2013), the existing papers on small states do not distinguish between these two types of government spending (Guy and Palgrave, 2012; Narita, 2014). Given their significant difference in the short and medium terms, it is crucial to separately estimate fiscal multipliers for these two types of government spending.

Notwithstanding the different methodologies, our results are qualitatively consistent with those of the literature, but they are quantitatively different (see table A5 in the online appendix). Our empirical results suggest that government consumption has an impact multiplier of around 0.3 and has a negligible medium-term impact on growth. Our GIMF model estimates a slightly larger impact multiplier of around 0.6, but it also finds a negligible medium-term impact on growth. On the other hand, our empirical and GIMF results both

23. Forni and Gambetti (2016) overcome the foresight problem in an SVAR framework by including forecast variables for U.S. data. However, none of the previous studies on small states addresses this issue.

suggest that government investment has a larger medium-term growth impact than government consumption, with fiscal multipliers at around 0.7 on average. These fiscal multipliers are in line with the results from the existing studies that estimate fiscal multipliers for small states.

Studies Based on Larger Samples

Some studies use broader samples that also include a number of small states. The IMF Regional Economic Outlook (IMF, 2018a), for instance, estimates fiscal multipliers for countries in Latin America and the Caribbean using a narrative approach, SVAR, and forecasting error methods; the study finds fiscal multipliers of between 0.5 and 1.1. For the narrative approach, the study uses annual data for the sample of fourteen Latin American and Caribbean countries between 1989 and 2016, combined with the fiscal consolidation episodes from David and Leigh (2018).[24] Their SVAR approach estimates fiscal multipliers country by country using quarterly data from eight Latin American and Caribbean countries. Finally, their forecast error approach uses annual data since 1990 for the sample of nineteen Latin American and Caribbean countries. They separate government consumption and government investment and estimate that the respective fiscal multipliers are 0.2 and 0.6 on impact and 0.5 and 1.1 after a year. However, the sample includes the larger countries in Latin America, which have higher GDP per capita than the small states included in our results. Nevertheless, the results suggest that fiscal multipliers are higher for government investment than for government consumption.

Batini and others (2014) review fiscal multipliers from the exiting literature, including for low-income and emerging economies. They show that fiscal multipliers are generally low for low-income and emerging economies, at around 0.2 to 1.3, with most panel studies finding multipliers around 0.2–0.5 on impact. Thus our empirical results are generally in line with the previous literature using different methodologies and different sets of countries across the world.

Concluding Remarks

This paper has offered a fresh look at fiscal multipliers for small states. We find that, in small states, short-term multipliers of government consumption (consumption in the empirical model) and investment are both around 0.4,

24. Argentina, Bolivia, Brazil, Chile, Colombia, Costa Rica, Dominican Republic, Ecuador, Guatemala, Jamaica, Mexico, Paraguay, Peru, and Uruguay.

on average, for the empirical and DSGE baseline results. In the medium-long term, government consumption (current primary spending in the empirical model) has a fiscal multiplier of about zero, but government investment has an average multiplier of around 0.7.

These results are consistent with the view that while government consumption can affect GDP in the short term, it does not affect potential GDP in small states. On the other hand, government investment affects both short-term and potential GDP in small states. Tax multipliers are found to be larger than government consumption multipliers but smaller than government investment multipliers. These multipliers are state dependent, and they are generally larger during recessions and consolidations than during booms and expansions. This asymmetry occurs because expansionary fiscal policy, especially in small states with high government debt, results in higher risk premiums (for example, on interest rates), which in turn dampen the multipliers.

This paper has several policy implications for small states. Governments that need to embark on a consolidation path are advised to design the composition in favor of cutting government consumption without cutting investment spending, as much as feasible. In fact, governments may find a consolidation plan to be growth friendly if, within the overall consolidation envelope, it includes an expansion of government investment. For governments that intend to embark on an expansion, the short-term benefits of current spending and investment are not materially different, while the medium-term benefits of investment are considerably larger than those of current spending.

There are several caveats to this study. First, the results may be affected by how government spending is financed. While the GIMF model assumes fiscal policy is financed by surplus/deficit, the empirical part does not differentiate between financing sources for government spending. For instance, despite the strong growth implication of public investment spending, using debt financing to increase public investment may not be a desirable policy tool as the return on public investment may not be sufficiently high to offset the interest on domestic and external loans. Second, given the annual data, our forecast error approach mitigates but does not fully solve the foresight (anticipation) problem and endogeneity issues. Our approach still leaves room for anticipation effects and endogeneity problems within the quarter. Moreover, our forecast error approach is as reliable as the forecasts we use. Finally, this study does not consider the political difficulty and possible distributional impact of different fiscal policy instruments.

References

Arora, Vivek, and Martin Cerisola. 2001. "How Does U.S. Monetary Policy Influence Sovereign Spreads in Emerging Markets?" *IMF Staff Papers* 48 (3): 474–98.

Auerbach, Alan J., and Yuriy Gorodnichenko. 2013. "Output Spillovers from Fiscal Policy." *American Economic Review* 103 (3): 141–46.

Batini, Nicoletta, Luc Eyraud, Lorenzo Forni, and Anke Weber. 2014. "Fiscal Multipliers: Size, Determinants, and Use in Macroeconomic Projections." Technical Notes and Manuals 14/04. Washington, DC: IMF.

Bernanke, Ben S., Mark Gertler, and Simon Gilchrist. 1999. "The Financial Accelerator in a Quantitative Business Cycle." In *Handbook of Macroeconomics*, vol. 1C, ed. Michael Woodford and John B. Taylor, chap. 21. Amsterdam: Elsevier Science.

Blanchard, Olivier J. 1985. "Debt, Deficits, and Finite Horizons." *Journal of Political Economy* 93 (2): 223–47.

Blanchard, Olivier J., and Roberto Perotti. 2002. "An Empirical Characterization of the Dynamic Effects of Changes in Government Spending and Taxes on Output." *Quarterly Journal of Economics* 117 (4): 1329–68.

Coenen, Günter, Christopher Erceg, Charles Freedman, and others. 2010. "Effects of Fiscal Stimulus in Structural Models." Working Paper 10/73. Washington, DC: IMF.

David, Antonio, and Daniel Leigh. 2018. "A New Action-Based Dataset of Fiscal Consolidation in Latin America and the Caribbean." Working Paper 18/94. Washington, DC: IMF.

Dodzin, Sergei, and Xuefei Bai. 2016. "Estimating Fiscal Multipliers Using a Simplified General Equilibrium Model of Small States, with Application to Kiribati and Palau." In *Resilience and Growth in the Small States of the Pacific*, ed. Hoe Ee Khor, Roger P. Kronenberg, and Patrizia Tumbarello, chap. 11. Washington, DC: IMF.

Forni, Mario, and Luca Gambetti. 2016. "Government Spending Shocks in Open Economy VARs." *Journal of International Economics* 99:68–84.

Furceri, Davide, and Bin Grace Li. 2017. "The Macroeconomic (and Distributional) Effects of Public Investment in Developing Economies." Working Paper 17/217. Washington, DC: IMF.

Furceri, Davide, Jun Ge, Prakash Loungani, and Giovanni Melina. 2018. "The Distributional Effects of Government Spending Shocks in Developing Economies." Working Paper 18/57. Washington, DC: IMF.

Gonzalez-Garcia, Jesus, Antonio Lemus, and Mico Mrkaic. 2013. "Fiscal Multipliers in the ECCU." Working Paper 13/17. Washington, DC: IMF.

Guy, Kester, and Anton Belgrave. 2012. "Fiscal Multiplier in Microstates: Evidence from the Caribbean." *International Advances in Economic Research* 18 (1): 74–86.

Ilzetzki, Ethan, Enrique G. Mendoza, and Carlos A. Végh. 2013. "How Big (Small?) Are Fiscal Multipliers?" *Journal of Monetary Economics* 60 (2): 239–54.

IMF (International Monetary Fund). 2015 *Annual Report on Exchange Arrangements and Exchange Restrictions (AREAER).* Washington, DC: IMF. Available online at www.elibrary-areaer.imf.org/Pages/Home.aspx.

———. 2017. "2017 Staff Guidance Note on the Fund's Engagement with Small Developing States." Washington, DC: IMF.

———. 2018a. "Fiscal Multipliers: How Will Consolidation Affect Latin America and Caribbean?" *Regional Economic Outlook: Western Hemisphere, April 2018*, chap. 4. Washington, DC: IMF.

———. 2018b. "World Economic Outlook, October 2018: Challenges to Steady Growth." Washington, DC: IMF.

Jordà, Òscar. 2005. "Estimation and Inference of Impulse Responses by Local Projections." *American Economic Review* 95 (1): 161–82.

Kumhof, Michael, and Douglas Laxton. 2007. "A Party without a Hangover? On the Effects of U.S. Government Deficits." Working Paper 07/202. Washington, DC: IMF.

———. 2009a. "Fiscal Deficits and Current Account Deficits." Working Paper 09/237. Washington, DC: IMF.

———. 2009b. "Simple, Implementable Fiscal Policy Rules." Working Paper 09/76. Washington, DC: IMF.

Kumhof, Michael, Douglas Laxton, Dirk Muir, and Susanna Mursula. 2010. "The Global Integrated Monetary Fiscal Model (GIMF): Theoretical Structure." Working Paper 10/34. Washington, DC: IMF.

Leigh, Daniel. 2008. "Achieving a Soft Landing: The Role of Fiscal Policy." Working Paper 08/69. Washington, DC: IMF.

Narita, Machiko. 2014. "Fiscal Multipliers in the Caribbean." In *Caribbean Renewal: Tackling Fiscal and Debt Challenge,* ed. Charles Amo-Yartey and Therese Turner-Jones, chap. 8. Washington, DC: IMF.

Ramey, Valerie A., and Sarah Zubairy. 2018. "Government Spending Multipliers in Good Times and in Bad: Evidence from U.S. Historical Data." *Journal of Political Economy* 126 (2): 850–901.

Romer Christina D., and David H. Romer. 2010. "The Macroeconomic Effects of Tax Changes: Estimates Based on a New Measure of Fiscal Shocks." *American Economic Review* 100 (3): 763–801.

Teulings, Coen N., and Nikolay Zubanov. 2014. "Is Economic Recovery a Myth? Robust Estimation of Impulse Responses." *Journal of Applied Econometrics* 29 (3): 497–514.

World Bank. 2016. *World Bank Group Engagement with Small States: Taking Stock.* Washington, DC: World Bank, pubdocs.worldbank.org/en/244361475521083722/Small-States-Stocktaking-2016.pdf.

ANTONIO C. DAVID
International Monetary Fund, United States

TAKUJI KOMATSUZAKI
International Monetary Fund, United States

SAMUEL PIENKNAGURA
International Monetary Fund, United States

The Macroeconomic and Socioeconomic Effects of Structural Reforms in Latin America and the Caribbean

ABSTRACT This paper estimates the macroeconomic effects of market-oriented reforms in Latin America and the Caribbean using the IMF Structural Reform database. We find that large changes in the reform index have positive effects on GDP that exceed 2 percent after five years. Furthermore, reforms boost employment, investment, exports, and imports and reduce export concentration, in addition to favoring tradable sectors. The evidence on the effects of reforms on business confidence is mixed, and the effects on total factor productivity are positive, but less precisely estimated. Nonetheless, our results also indicate that the effects of reforms have not been uniform across different segments of the population. Our results are robust to the use of an instrumental variables approach that exploits regional waves of reform to deal with endogeneity concerns. These findings bring to the forefront the need to consider accompanying policies to ensure that reforms promote inclusive growth.

JEL Codes: E20, O11, O40

Keywords: Structural reforms, Latin America, macroeconomic effects

Economic growth in Latin America and the Caribbean has been sluggish for a prolonged period. Labor and total factor productivity (TFP) growth have lagged those of other emerging markets and developing economies. This situation is, in part, linked to significant structural constraints, including inadequate infrastructure, high levels of informality, low levels of human capital, and weak governance (Bakker and others, 2020).

ACKNOWLEDGMENTS We thank the editor, Bernardo Guimarães, and two anonymous referees for excellent suggestions. We are grateful to Chris Papageorgiou for kindly sharing the data set on structural reforms. We are also grateful to Romain Duval, Dmitry Gershenson, Luca Ricci, Jorge Roldos, and seminar participants at the IMF for very helpful comments. Genevieve Lindow provided outstanding research assistance. The views expressed in this paper are those of the authors and do not necessarily represent the views of the International Monetary Fund (IMF), its Executive Board, or IMF management.

To overcome stagnation, countries in the region have undertaken important efforts to liberalize key markets, particularly in the 1990s and 2000s. These efforts were followed by reform fatigue and, in some cases, reversals. Could this pattern be partly grounded in a perception by policymakers and the general public that reforms failed to deliver? Does the empirical evidence validate such perceptions regarding disappointing gains from past reforms? Or have reforms delivered positive outcomes, but not for all segments of the population?

This paper addresses these questions by estimating the effects of specific reforms—namely, trade, product market, labor market, and domestic financial liberalization—on key macroeconomic and social variables. A significant contribution of the paper to the literature on structural reforms in emerging market economies is to extend the analysis beyond the usual aggregates, such as GDP, and zoom in on key transmission channels through which reforms affect macroeconomic outcomes over the short to medium term, such as total investment, foreign direct investment, informality, business confidence, and sectoral effects. Moreover, the paper also studies the potential collateral damage of reforms, given that reforms with significant negative effects on inequality and poverty are unlikely to be sustainable.

Using the International Monetary Fund's structural reform database as first employed in Alesina and others (2020), we find that large changes in the index (toward liberalization) have positive effects on GDP and employment in Latin America and the Caribbean, which reach 2.4 percent and 1.7 percent after five years, respectively. Market-oriented reforms also increase TFP, but their effects are more imprecisely measured. Nonetheless, the results also suggest that reforms have had economically small but statistically significant adverse effects on inequality and poverty.

The positive effects of reforms on aggregate growth appear to operate through specific channels, namely, higher investment and de facto openness. Reforms boost investment, real exports, and real imports and reduce export concentration, in addition to favoring tradable sectors. The evidence on the effects of reforms on business confidence is more mixed, and there is no evidence that reforms significantly affect informality. There is also evidence of complementarities between reforms.

Ensuring that these findings are indeed caused by market-oriented reforms requires careful consideration of potential endogeneity issues. First, market-oriented liberalization is not exogenously given to countries; countries self-select to pursue reforms. For example, countries may choose to take reform actions in response to low growth and employment. Alternatively, countries may have inherent differences that affect both the decision to pursue market-oriented liberalization and growth. Second, reform efforts may coincide with periods of commodity booms and busts, which is an important concern since our sample is skewed toward commodity exporters. Our baseline specifications partially deal with these potential sources of endogeneity by including lags of the commodity terms-of-trade index, past growth, and country fixed effects as control variables. However, other concerns remain, since the liberalization decision may be correlated with time-variant unobservable variables such as expectations of future growth, and countries that decide to liberalize may have higher growth prospects (Buera, Monge-Naranjo, and Primiceri, 2011).

To address these remaining concerns, we also implement an instrumental variables (IV) approach that exploits regional waves of reform. More precisely, we construct a distance-weighted index of reforms in nearby countries and use changes in the index as an instrument. A similar IV strategy has been used to study the causal effects of democratization on growth (Acemoglu and others, 2019) and the impact of fiscal austerity on social unrest (Ponticelli and Voth, 2020). In the specific case of episodes of liberalization and reform reversal, the exercise is grounded in the theoretical findings of Buera, Monge-Naranjo, and Primiceri (2011). Reassuringly, the findings from the IV approach corroborate our baseline ordinary least squares (OLS) results.

This paper is related to a long-standing literature on the state of the structural reform agenda in developing countries and its effects on growth (see Zettelmeyer, 2006, for a summary of the effects of reforms in Latin America and the Caribbean). It is closely linked to IMF (2019) and Alesina and others (2020), which study the effects of structural reforms on growth and informality in a large set of countries. We expand their analysis by zooming in on the channels through which reforms may affect growth, and focus exclusively on Latin America and the Caribbean.

As in Lora (2012) and IMF (2019), the analysis here unbundles the state of the reform agenda along different dimensions. Doing so allows us to study the differential effects of specific reform areas. In this regard, the paper is also related to Biljanovska and Sandri (2018), who study the effects of different reforms on TFP growth in Brazil. This paper broadens the focus to a larger set of countries and focuses on the dynamic response of macroeconomic variables following reform episodes. The effects of reforms on economic development are also studied in Bergoeing and others (2001), who compare the economic development path of Chile and Mexico in the 1980s and 1990s and argue that policy reforms implemented in Chile fostered faster productivity growth. The findings in Billmeier and Nannicini (2013) also provide support to the link between reforms (liberalization) and growth, especially during the first wave of reforms in the 1980s. In addition, Prati, Onorato, and Papageorgiou (2013) find that while reforms are positively associated with higher growth on average, this link is highly heterogeneous and seems to be influenced by a country's institutions and distance from the technology frontier.

A related literature attempts to explain the drivers of reforms rather than their economic effects, which remains our main focus in this paper (Buera, Monge-Naranjo, and Primiceri, 2011; Dias Da Silva, Givone, and Sondermann, 2017; Prati, Onorato, and Papageorgiou, 2013; Duval, Furceri, and Miethe, 2021). The seminal work by Buera, Monge-Naranjo, and Primiceri (2011) explores how a country's own and its neighbors' past experiences influence policy choices through their effect on policymakers' beliefs. They find that the evolution of beliefs about the relative desirability of free markets can be a major driving force behind regime transitions (between market orientation and state intervention). Overall, from an empirical perspective, papers in the literature on the drivers of reforms also tend to find some evidence that crises are associated with subsequent reform upticks and that there is reform convergence (such that countries with tighter regulation are more prone to liberalize).

Our paper is structured as follows. The next section presents some stylized facts about reforms in Latin America and the Caribbean since the 1970s, including a discussion of public opinion surveys gauging support for reforms in the region. We then quantify the effects of reforms on GDP, employment, and TFP and assess whether the effects of reforms vary with the state of the economic cycle and whether there are complementarities between reforms. This section also looks at a number of transmission channels that might mediate the effects of reforms on GDP—such as total investment, foreign direct investment, informality, business confidence, external trade, and the shares of different sectors in the economy—and considers the effects of reforms on poverty and inequality. Subsequently we present the results of the IV strategy exploiting regional waves of reforms. The final section concludes.

Structural Reform Efforts in Latin America and the Caribbean Since the 1970s

It is difficult to measure structural reform efforts consistently across countries and time. This paper follows the approach of IMF (2019) and Alesina and others (2020) by focusing on some specific aspects of reforms that aim to liberalize certain markets. The analysis is mostly based on the IMF Structural Reform Database, which was updated up to 2018 for the trade liberalization component. The data set covers reforms implemented in ninety countries over the period 1973–2014, at an annual frequency.[1] Higher values of the index point to more liberalized and better regulated areas, but there are also several instances of reform reversals in the database.

Using these data, we analyze reforms implemented in four broad areas: (1) domestic finance, which includes six dimensions of domestic finance regulation (credit controls, interest rate controls, entry barriers, supervision, privatization, and security markets development); (2) trade, based on average tariff levels; (3) product market, which considers liberalization and regulation in two network sectors (telecommunications and electricity) covering three broad areas (privatization, entry barriers, and supervision and regulation); and (4) labor market, which provides a measure of employment protection legislation covering four areas (procedural requirements, firing costs, valid grounds for dismissal, and redress measures). IMF (2019) provides a description of the indicators and criteria used to build the reform indexes along these four dimensions.

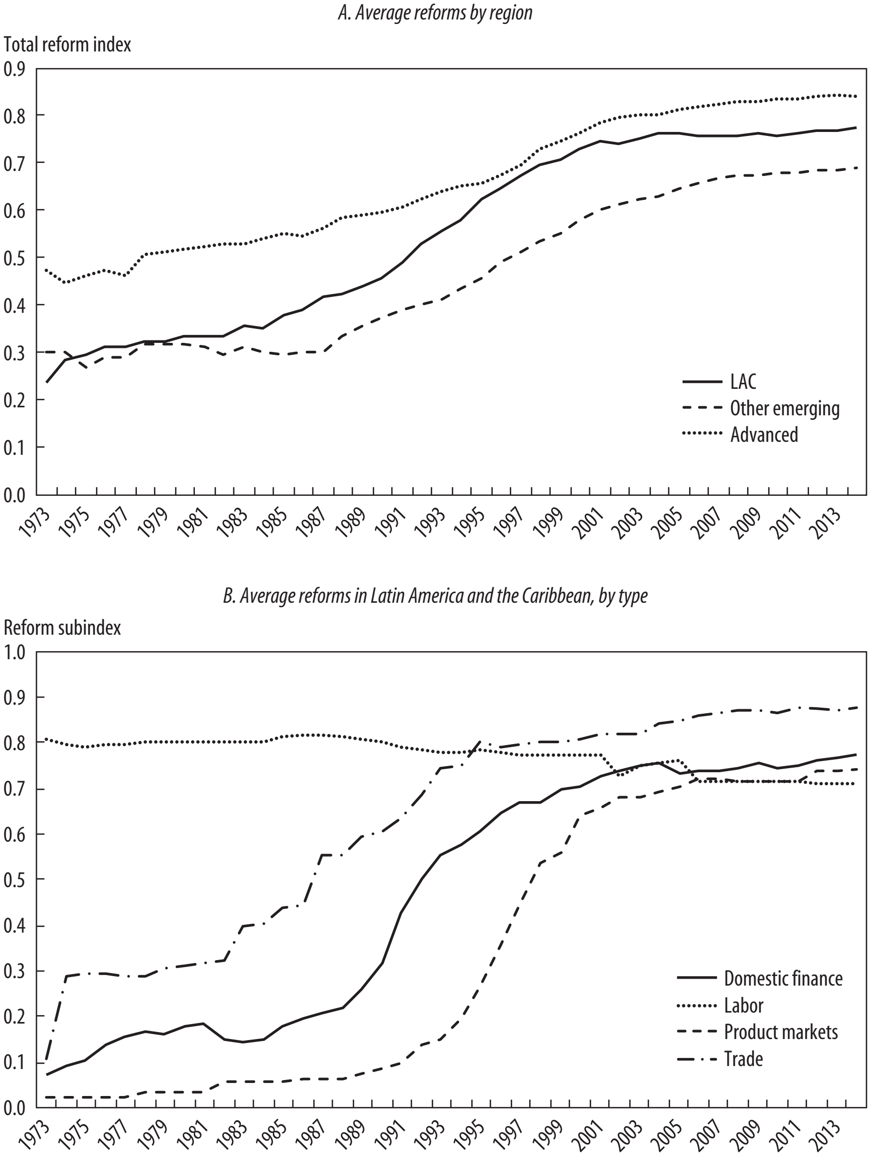
Figure 1 depicts an overall index of reforms in the region as the simple average of the four dimensions outlined above, normalized to take a value between zero and one, with one being the most liberalized and better regulated. Data show that the typical country in the region undertook substantial reforms in the 1990s and early 2000s, but reform impetus has stalled somewhat in more recent periods. Despite notable progress, the region lags advanced economies on the overall index and on some reform dimensions. With respect to specific reform areas, on average, countries in the region have taken steps to liberalize trade, product markets, and domestic finance over the 1990s and 2000s, while reforms to employment protection legislation have been less frequent.[2]

Regional averages mask significant heterogeneity across countries. As illustrated in figure 2, progress in terms of specific reform areas varies substantially

1. The sample includes sixty-eight emerging and developing economies, of which seventeen are in Latin America and the Caribbean.

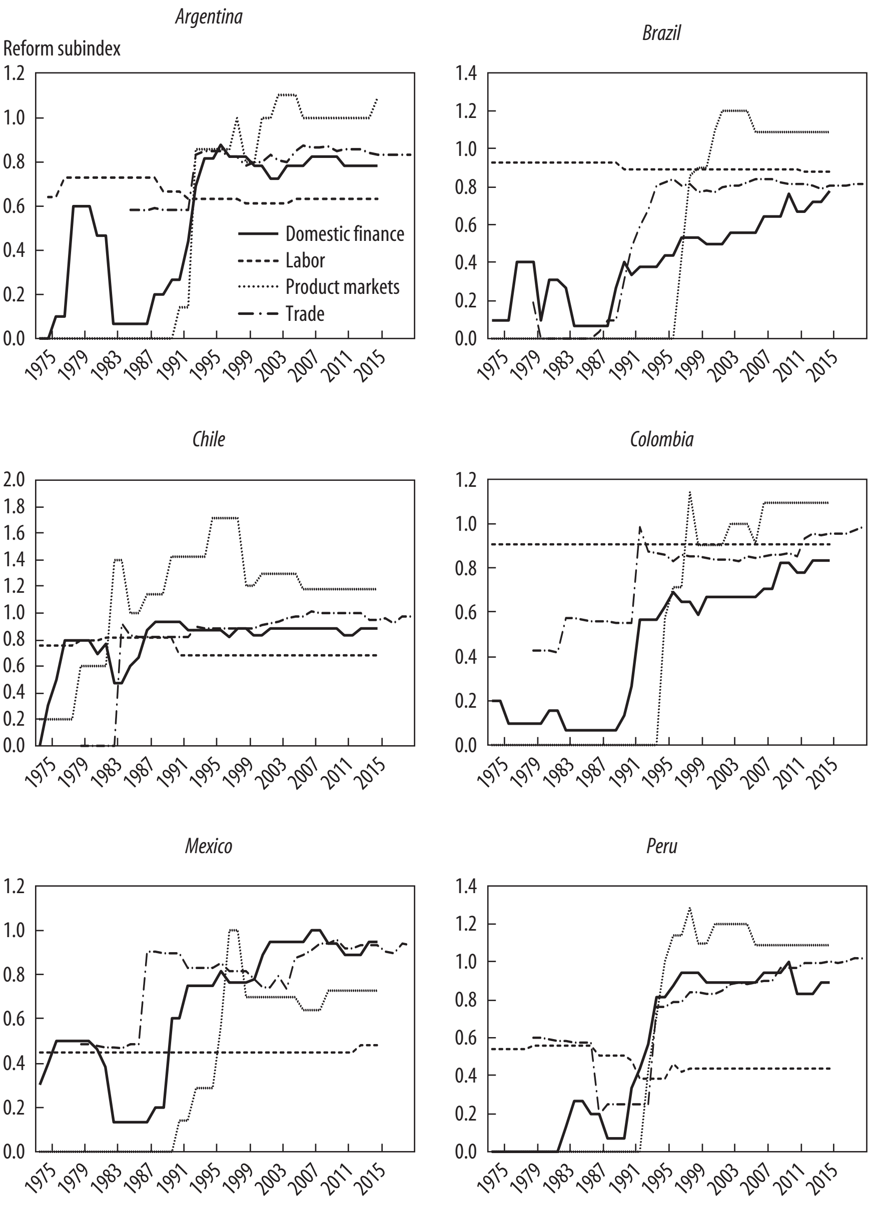
2. As explained in IMF (2019), because of the nature of the indicators, one cannot directly compare a country's regulatory stance across different areas. All comparisons need to be made relative to other countries. Thus increases in the indexes for the different areas point to steps taken toward liberalization, but it is not possible to claim, for example, that trade is more liberalized than labor markets in a given country just by directly comparing the levels of these indicators. For this reason, we turn to ratios relative to the United States next.

FIGURE 1. Trends in the Structural Reform Index



Source: Authors' calculations, based on IMF data.
Note: LAC, Latin America and Caribbean.

FIGURE 2. Ratios Relative to the United States for Different Reform Areas



Source: Authors' calculations, based on IMF data.

across some of the largest economies in Latin America and the Caribbean. The figure depicts the ratios of specific reform indexes in a given country relative to the United States, hence indicating whether the country is more or less liberalized in one particular area. For example, Brazil still has ground to cover in terms of trade and domestic financial liberalization, while Mexico lags in the areas of labor and product market reforms. Moreover, several countries still seem to have particularly stringent employment protection legislation, including Argentina, Chile, Mexico, and Peru.

What drives the implementation of reforms? Buera, Monge-Naranjo, and Primiceri (2011) use a learning model fitted to a panel of countries over the period 1950–2001 to show that the evolution of beliefs about the relative desirability of free markets can be a major driving force behind transitions between market-oriented regimes and regimes based on state intervention. In their model, policymakers have initial priors about the relative growth prospects of different regimes and use Bayes' theorem to update these priors with the arrival of new information from all countries in the world. A country will decide to pursue market-oriented policies if the perceived net impact of these policies on GDP growth exceeds their political cost.

Dias Da Silva, Givone, and Sondermann (2017) find that reforms are more likely during deep recessions and when the unemployment rate is high, based on a sample of forty countries in the Organization for Economic Cooperation and Development (OECD) and the European Union (EU). Distance from the frontier is also an important empirical determinant of reforms. The presence of an IMF-supported program or other forms of external conditionality also facilitates reforms, but there is no clear link between fiscal policies and reforms. Prati, Onorato, and Papageorgiou (2013) also find some evidence that severe growth downturns are associated with subsequent reform upticks, based on a larger sample of countries.

These findings are broadly confirmed by Duval, Furceri, and Miethe (2021) for product and labor market reforms in a sample of advanced economies, using Bayesian model averaging techniques. They find evidence to support the hypothesis that economic crises induce reforms and also conclude that there is reform convergence (that is, countries with tighter regulation are more prone to liberalize). Reforms are more likely when other countries also undertake them and when there is external pressure to implement them (such as during IMF-supported programs).

In contrast, Ciminelli and others (2019), based on a broad sample of countries, find that reforms are often reversed during periods of low growth. The effects of economic downturns on reforms also tend to vary depending on the reform area (IMF, 2019). Recessions foster trade, labor market, and domestic financial liberalization, but banking crises are linked to reversals in domestic and external financial liberalization.

The impetus for reform has declined in several countries in Latin America and the Caribbean since the 2000s. To explore whether this trend reflects a perception by the general public and policymakers that reforms failed to deliver, we follow Biljanovska and Sandri (2018) and use information from the Latinobarómetro public opinion surveys over several years to gauge support for reforms in the region. Overall support for reforms is proxied by the share of survey respondents who express support for the market system by indicating whether they agree or strongly agree with the statement that the market economy is the only system with which the country can become developed. Figure 3 shows that there is broad support for market liberalization across countries in the region (panel A). In several countries, however, as many as a quarter to a third of respondents expressed skepticism of reforms, as proxied by the share of respondents who disagree or strongly disagree with the above statement (panel B). With regard to specific reform areas, the share of respondents supporting trade liberalization is generally low across the region, especially in Central America and Mexico (panel C). Support for finance and product market reforms (proxied by the share of respondents supporting innovation and productivity, following Biljanovska and Sandri, 2018) is higher than support for trade integration across the region, but it is particularly high in South American countries and Costa Rica.

Thus, while there is, in general, broad support for reforms across countries in the region, opinion surveys also suggest that a significant share of the population remains skeptical regarding the benefit of reforms, particularly in areas such as trade liberalization. In that context, an empirical assessment of the economic effects of reforms becomes particularly relevant. We turn to this issue in the next section.

Quantifying the Effects of Structural Reforms

This section studies the effects of reforms on real GDP, employment, and total factor productivity (TFP) over the medium term for the seventeen Latin American and Caribbean countries in the data set using the local projection method. This procedure does not constrain the shape of the impulse response

FIGURE 3. Support for Reforms

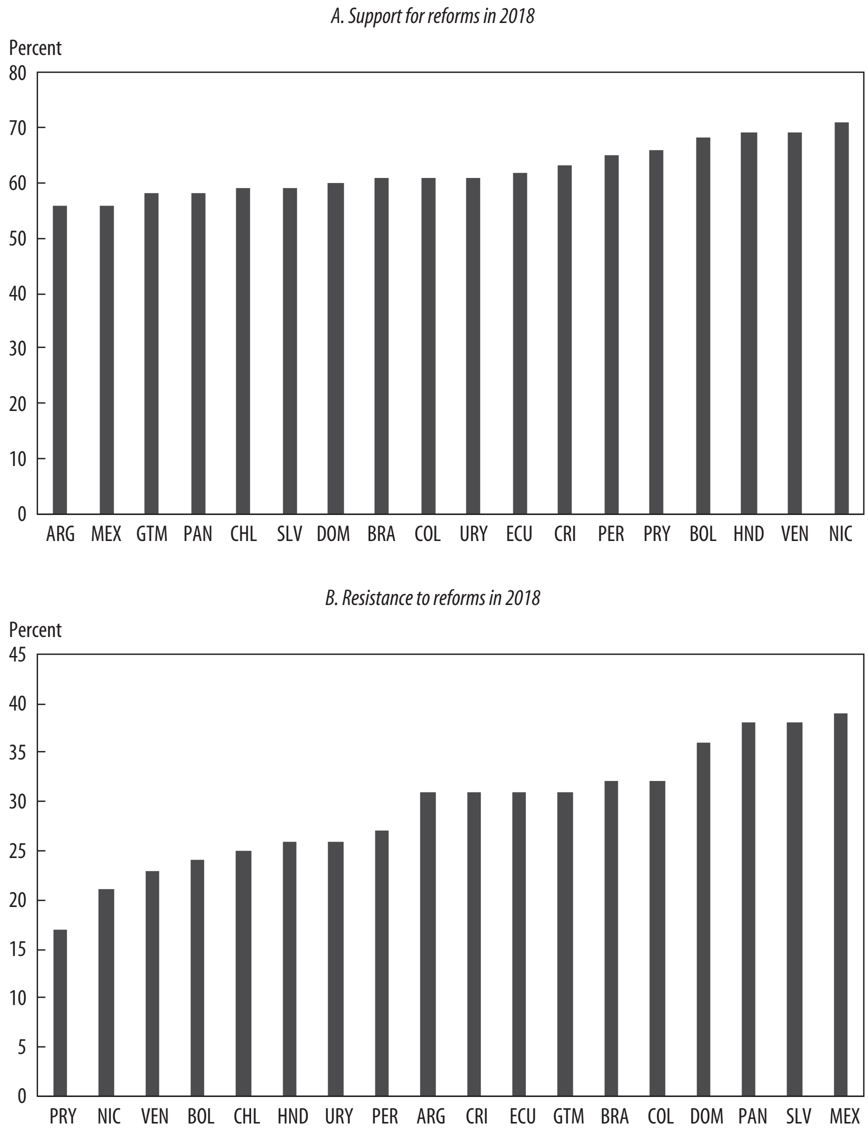
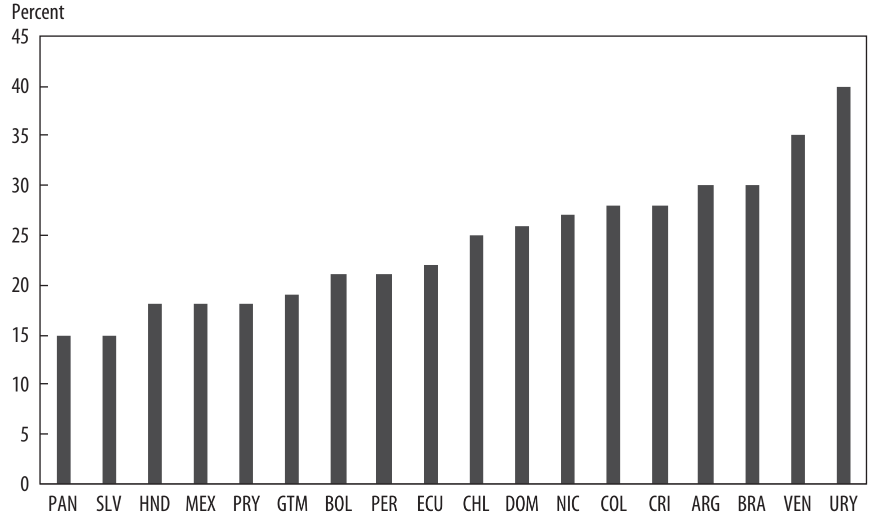
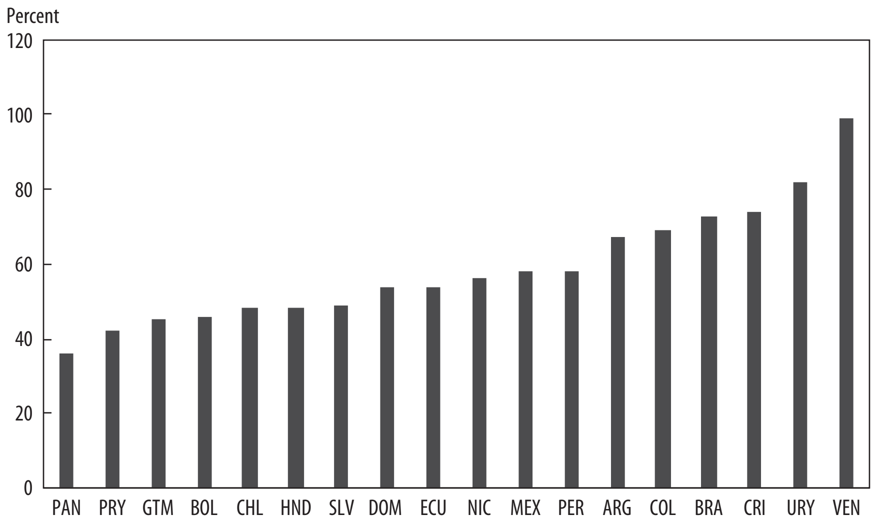


FIGURE 3. Support for Reforms (*Continued*)

C. Support for trade liberalization in 2017



D. Support for finance and product market reforms in 2017



Source: Authors' calculations, based on Latinobarómetro (several years).

functions and is therefore less sensitive to misspecification than estimates of vector autoregression (VAR) models (Jordà and Taylor, 2016). The benchmark specification at an annual frequency is as follows:

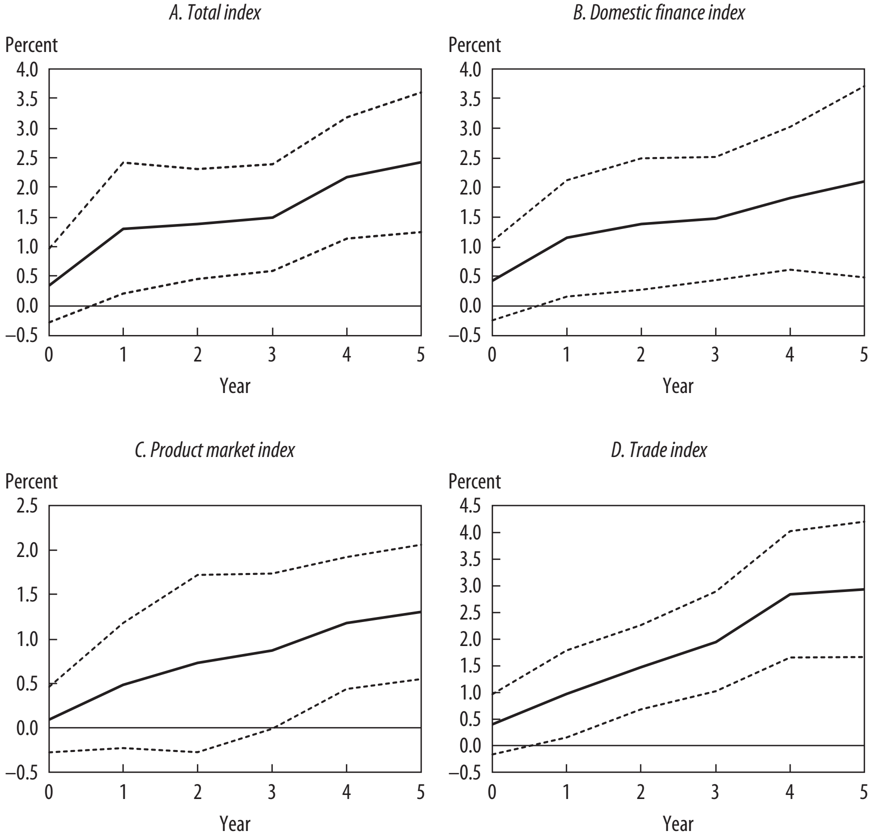
$$(1) \qquad y_{i,t+h} - y_{i,t-1} = \alpha_i^h + \gamma_t^h + \beta^h \Delta SR_{i,t} + \delta \mathbf{X}_{i,t} + \varepsilon_{i,t+h},$$

where y denotes the variable of interest (real GDP, employment, or TFP in this section, while subsequent sections will focus on other dependent variables, such as investment, informality, and inequality); $\Delta SR_{i,t}$ denotes the change in the structural reform index; and h denotes the time horizons considered. The vector $\mathbf{X}_{i,t}$ denotes a set of control variables, which includes lagged values of the dependent variable and of the reform index, as well as changes in the commodity terms-of-trade index constructed by Gruss and Kebhaj (2019) and its lags, which were included because terms of trade are an important driver of the business cycle in emerging market economies (Fernández, González, and Rodríguez, 2018). The specification also includes time (γ_t^h) and country (α_i^h) fixed effects to capture common shocks and time-invariant country features, respectively.[3] We present impulse responses for large changes in the reform indexes (two standard deviations) in the figures below. The appendix provides definitions and sources for the main variables used in the analysis.

While the local projection method provides a flexible framework to estimate the dynamic effects of reforms, the approach by itself does not solve endogeneity issues arising from reverse causality and omitted variables. In that context, our main identification assumption in the paper relies on a timing restriction that reforms take time to implement and typically are not caused by movements in the dependent variable of interest within the same year. Later in the paper, we address endogeneity concerns by implementing an instrumental variables approach that exploits liberalization episodes in nearby countries as a potential source of exogenous variation. Regarding omitted variables, as discussed above, our regressions include country and year fixed effects that allow us to control for time-invariant country-specific factors and common shocks (across countries in the sample), respectively. In addition, omitted variables bias is further attenuated by the inclusion in all regressions of two lags of the dependent variable, as well as lags of the reform indexes and the commodity terms-of-trade variable.

3. For specifications that consider specific reforms (rather than the total reform index), we also add as controls lagged values of the changes in the other reform indexes to control for complementarities.

FIGURE 4. Effects of Reforms on GDP



Source: Authors' calculations, based on IMF data.
Note: The dashed lines show the 90 percent confidence interval for Driscoll-Kraay standard errors.

Figure 4 shows the effects of large changes in the reform index on real GDP. Reforms in the region have positive effects on GDP that reach 2.4 percent after five years (panel A). This estimated magnitude of the effects of reforms is in line with the average findings of the IMF (2019) for a broader set of emerging markets and developing economies.

We also consider specifications for reforms in specific areas, in which, in addition to the control variables discussed in equation 1, we add lagged changes of the other reform indicators to control for possible complementarities across reforms. Domestic finance reforms present a similar impulse response to the overall reform index, while product market reforms have positive effects on GDP that tend to take longer to materialize and become statistically significant only after two years. The effects of trade reforms on GDP are somewhat larger than the ones obtained for the overall index, reaching close to 3 percent after five years. The effects of labor market reforms on GDP for the sample of countries are not statistically significant and are not reported, to save space.[4]

Figure 5 presents the effects of reforms on employment (defined as the log of employment in thousands of people, from the World Development Indicators database). Reforms also tend to boost employment, with large changes in the total reform index being associated with increases in employment of 1.7 percent after five years, even if such increases tend to take time to materialize. Product market reforms, in particular, are linked to statistically and economically significant increases in employment one year after implementation.

We now turn to evidence on the effects of reforms on TFP. We take the TFP measure directly from Feenstra, Inklaar, and Timmer (2015). The impulse responses depicted in figure 6 show that reforms have a positive effect on TFP of about 1 percent, which is imprecisely measured for the total reform index (the confidence interval is wide) and only marginally significant at the 10 percent level at the five-year horizon. Nevertheless, when we focus on the trade reform index, the positive effects are statistically significant after two years, reaching about 1 percent after five years.

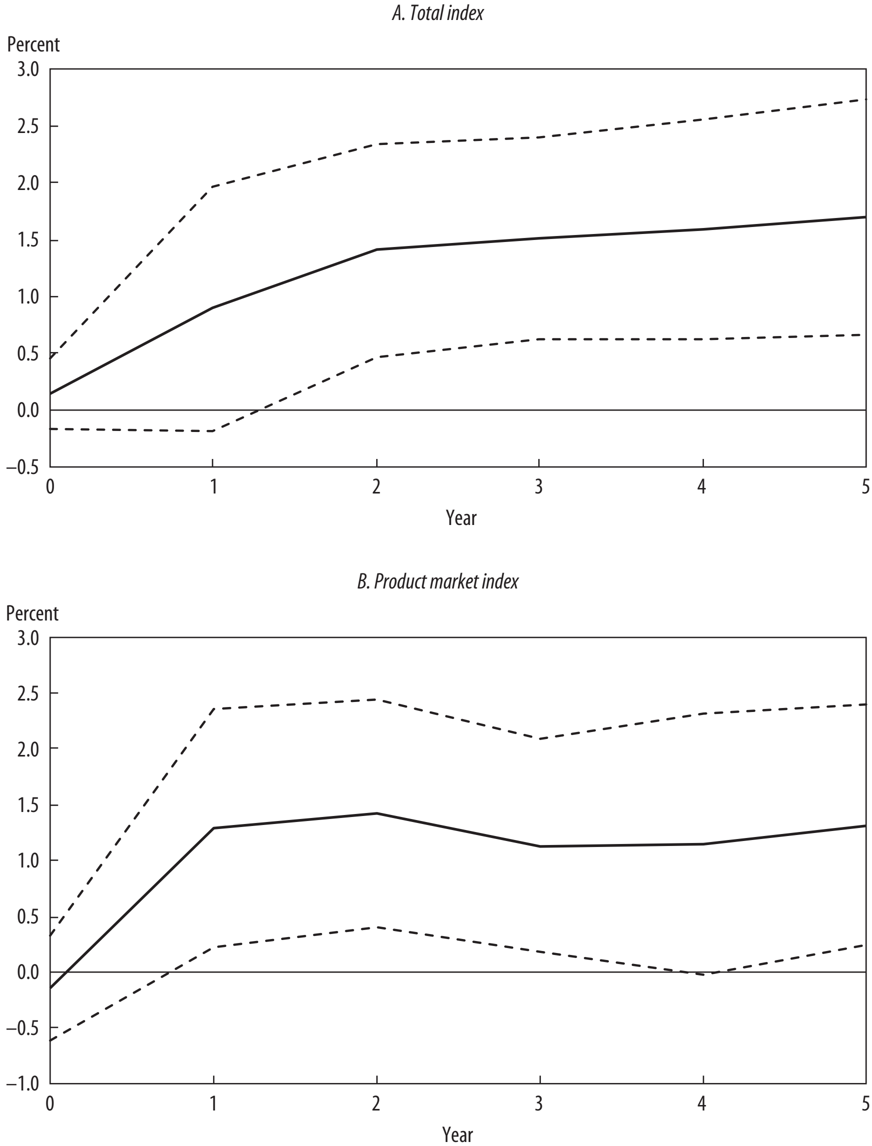
Overall, we find that reforms that move toward greater liberalization can have positive effects on output and employment for countries in Latin America and the Caribbean, but these benefits tend to take time to materialize. There is also evidence of positive effects in terms of TFP, but these effects are less precisely estimated.

Do Initial Conditions Matter?

We examine whether the baseline results change depending on conditions prevailing at the time of reform implementation. One of the main advantages of the local projection method is its flexibility in dealing with nonlinearities and state dependency. The typical state-dependent specification will take the

4. There are only twenty-five instances of nonzero changes for the labor reform indicator in Latin America and the Caribbean.

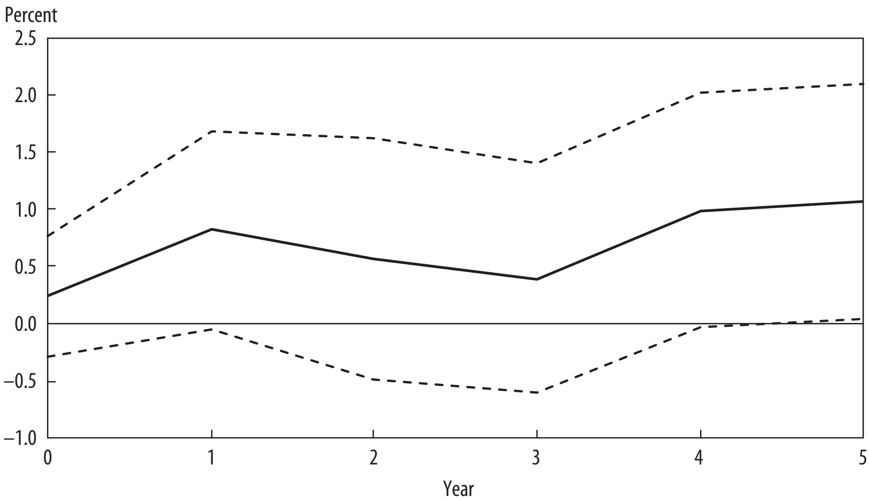
FIGURE 5. Effects of Reforms on Employment



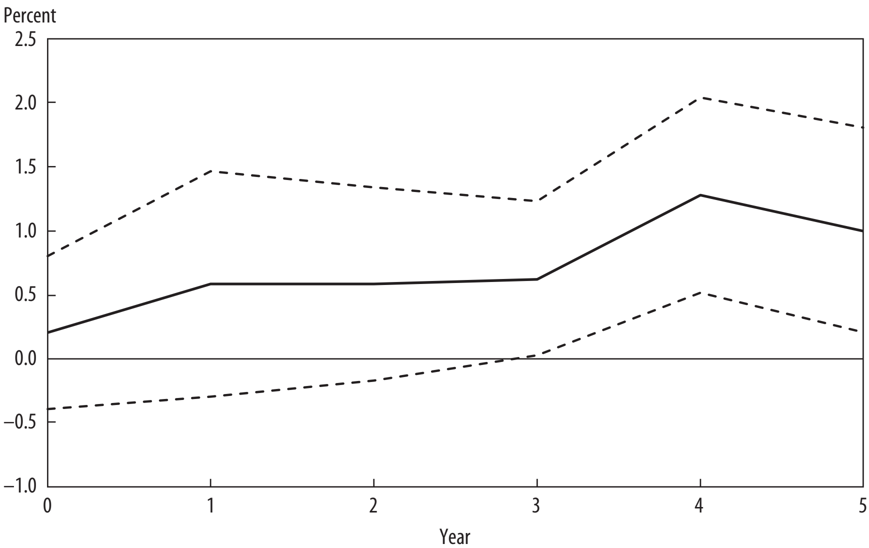
Source: Authors' calculations, based on World Development Indicators data.
Note: The dashed lines show the 90 percent confidence interval for Driscoll-Kraay standard errors.

FIGURE 6. Effects of Reforms on Total Factor Productivity

A. Total index



B. Trade index



Source: Authors' calculations, based on IMF data.
Note: The dashed lines show the 90 percent confidence interval for Driscoll-Kraay standard errors.

following form, with $S_{i,t-1}$ being an indicator variable taking the value of zero or one depending on the state dependency being considered:

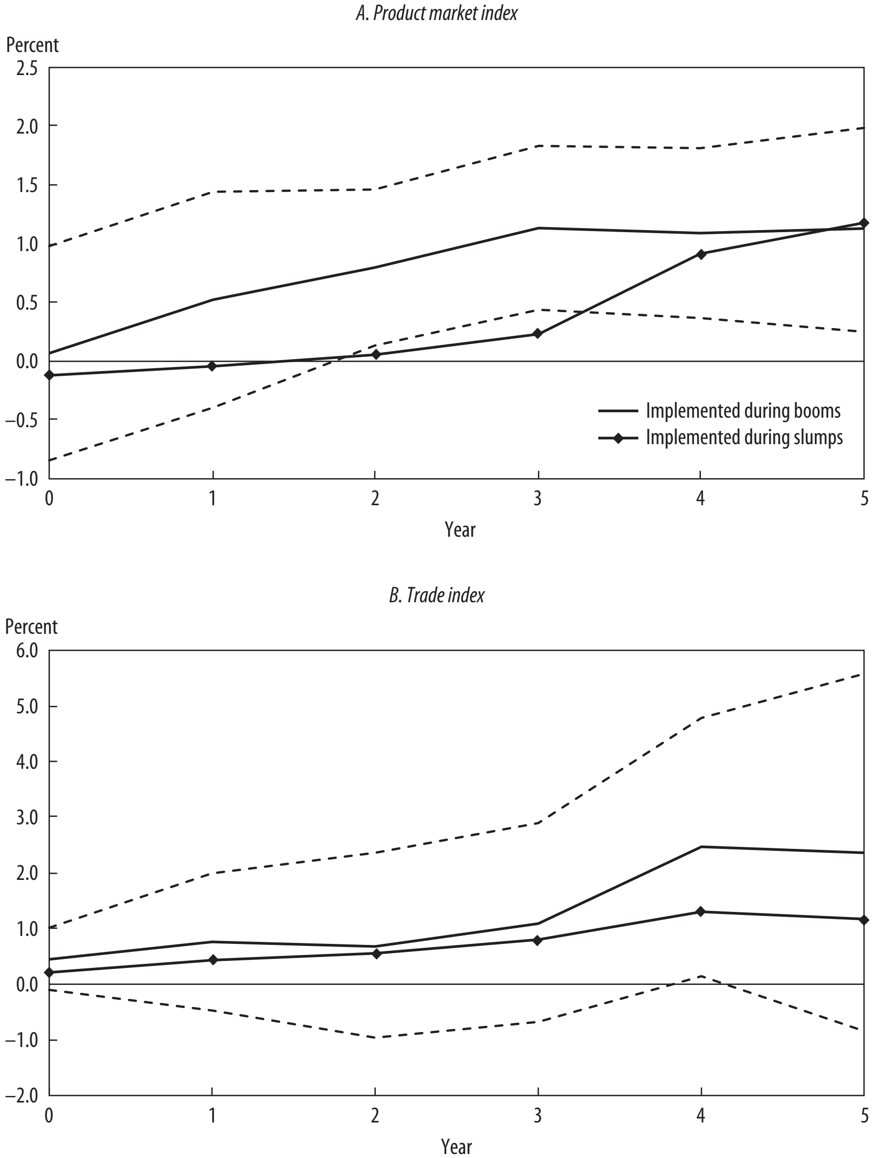
$$(2) \quad y_{i,t+h} - y_{i,t-1} = S_{i,t-1}\left(\alpha^h_{high,i} + \gamma^h_{high,t} + \beta^h_{high}\Delta SR_{i,t} + \delta_{high}\mathbf{X}_{it}\right) + \left(1 - S_{i,t-1}\right)\left(\alpha^h_{low,i} + \gamma^h_{low,t} + \beta^h_{low}\Delta SR_{i,t} + \delta_{low}\mathbf{X}_{it}\right) + \varepsilon_{i,t+h}.$$

We begin by analyzing whether the effects of reforms change depending on whether they were implemented in periods of economic expansion (boom) or contraction (slump). These periods were identified such that boom periods are years in which the output gap is positive (above trend GDP, which is estimated using the Hodrick-Prescott filter) and slump periods are years in which the output gap is negative.

For our sample of Latin American and Caribbean economies, the effects do not vary much according to the state of the economy for the total reform index, but the results do differ for some specific reforms. In particular, for product market and trade reforms, the effects on GDP are somewhat larger when they are implemented in boom times (figure 7). In the case of product market reforms, the difference disappears at the end of the five-year horizon. This is somewhat different from the general findings of IMF (2019) for a large sample of emerging and developing economies. That study found a marked contrast in the effects of reforms on GDP when the reforms were implemented in booms rather than recessions.

We also use the state-dependent specification outlined above to explore the role of possible complementarities between the different reform areas. To do so, we condition the impulse responses for a given reform (say, domestic finance) on whether the level of the reform index for other areas (trade, product market, and labor) is above or below the median for our sample of countries in the year before the implementation of the reform of interest. Figure 8 depicts the results of this exercise for the effects of domestic finance reforms on GDP, conditioned on the level of trade liberalization and on the level of employment protection liberalization. The impulse responses in panel A show that domestic finance reforms in the region have a positive effect on GDP, even when they are implemented at times when the economy is relatively more closed (that is, conditioning on lower levels of the trade liberalization index). Moreover, domestic finance reforms also have a positive effect on GDP even when labor markets are relatively rigid (panel B). Taken together, these results indicate that the positive payoffs of domestic finance reforms are

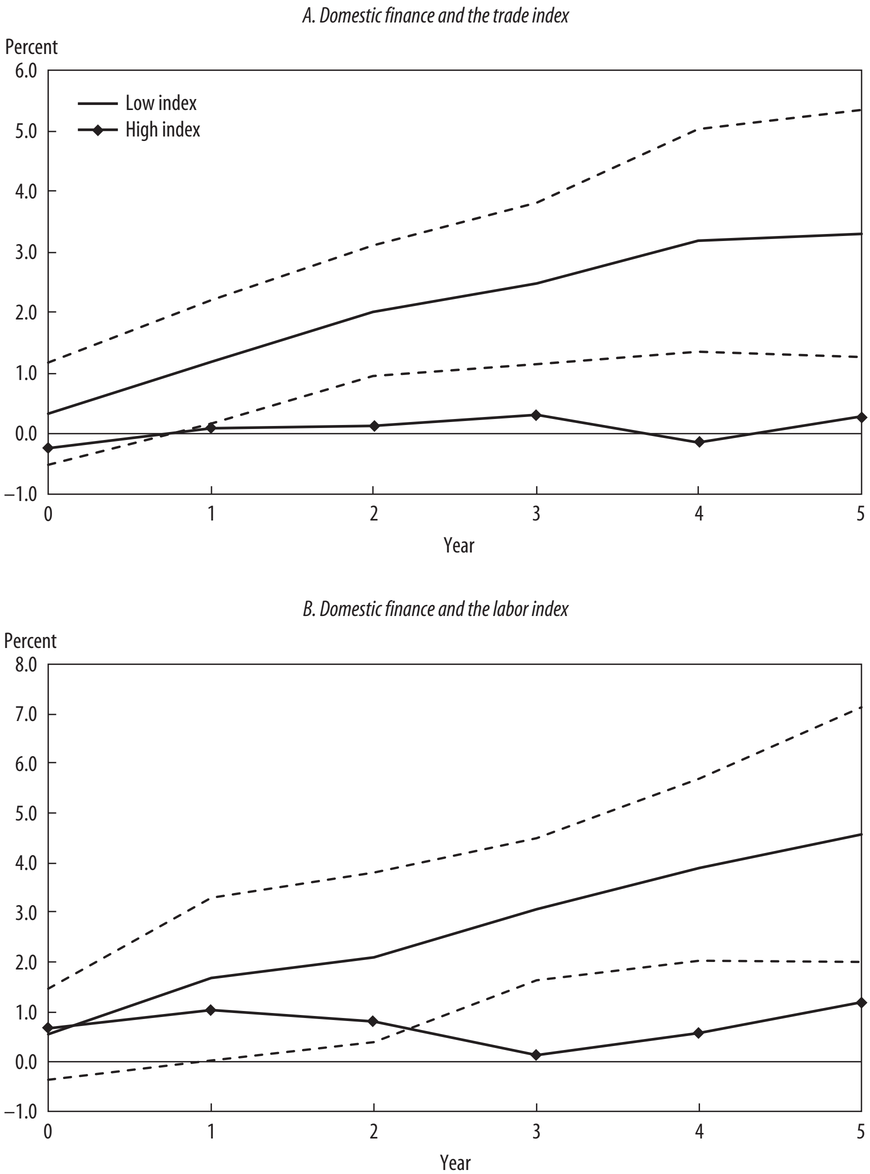
FIGURE 7. Effects of Reforms on GDP Depending on State of the Economy



Source: Authors' calculations, based on IMF data.

Note: The dashed lines show the 90 percent confidence interval for Driscoll-Kraay standard errors for reforms implemented during boom periods.

FIGURE 8. Effects of Domestic Finance Reforms on GDP, Conditioned on the Level of Liberalization



Source: Authors' calculations, based on IMF data.
Note: The dashed lines show the 90 percent confidence interval for Driscoll-Kraay standard errors for the low index.

not precluded by the presence of significant distortions in other areas (namely, trade and labor markets).

Similarly, as illustrated in figure 9, the effects of trade liberalization on GDP are positive even when employment protection legislation is relatively more rigid. The effects of product market reforms are also positive and significant even when implemented in periods of rigid employment protection legislation.

Inspecting the Mechanisms: Investment, FDI, Informality, and Confidence

This section considers the empirical effects of reforms on investment, foreign direct investment (FDI), informality, and business confidence indicators using a similar specification to equation 1. The purpose is to identify the mechanisms through which reforms affect GDP and employment. These channels have not received much attention in the literature.

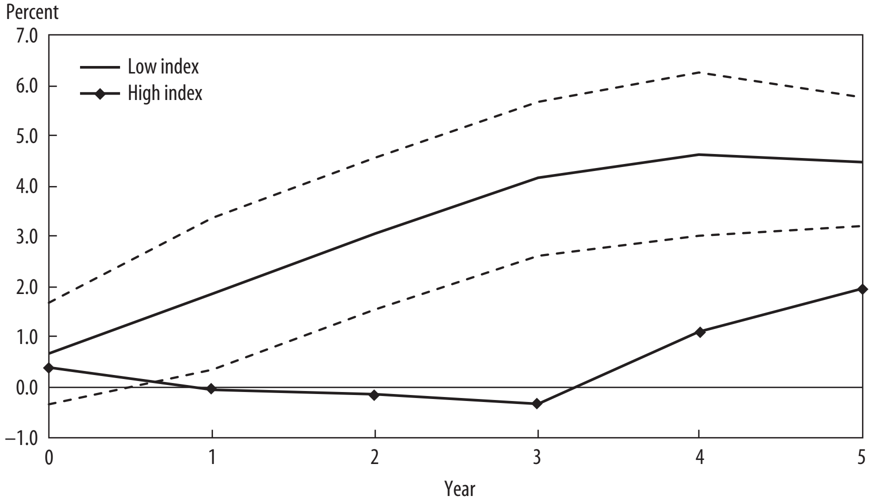
Figure 10 presents the results of the effects of reforms on total investment (in log real terms) and FDI (as a share of GDP). Large changes in the total structural reform index increase total investment by over 3.6 percent in a five-year period. The effects of domestic finance reforms on investment are particularly apparent, leading to increases that exceed 4 percent after two years, but over the medium term the confidence interval widens and the effects are no longer statistically significant in this case.

Reforms also boost FDI, although the effects tend to be economically small and only marginally significant from a statistical point of view. In the case of product market reforms, the effects are statistically significant but remain economically small: a two-standard-deviation change in the reform index is linked to an increase in FDI of a little more than 0.2 percent of GDP (and a peak increase of less than 0.4 percent of GDP).

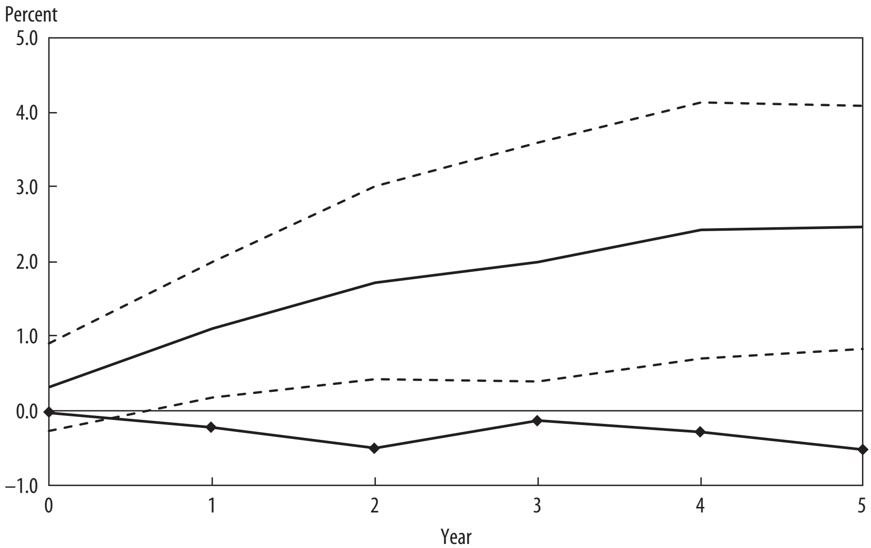
Latin American and Caribbean economies are marked by high levels of informality, which has important macroeconomic implications, including with regard to the adjustment to shocks (David, Roldos, and Pienknagura, 2020). Therefore, the effects of reforms on informality are particularly important from a policy perspective in the region. Figure 11 depicts how the informality rate, defined as the share of active workers not contributing to social security, responds to changes in the structural reform index. Changes in the total reform index are associated with a decrease in informality, but it is not statistically significant over the medium term. Nevertheless, when we consider product market reforms more specifically, the effects become statistically significant, though still economically small, with large reforms reducing the informality

FIGURE 9. Effects of Trade and Product Market Reforms on GDP, Conditioned on Labor Market Liberalization

A. Trade and the labor index

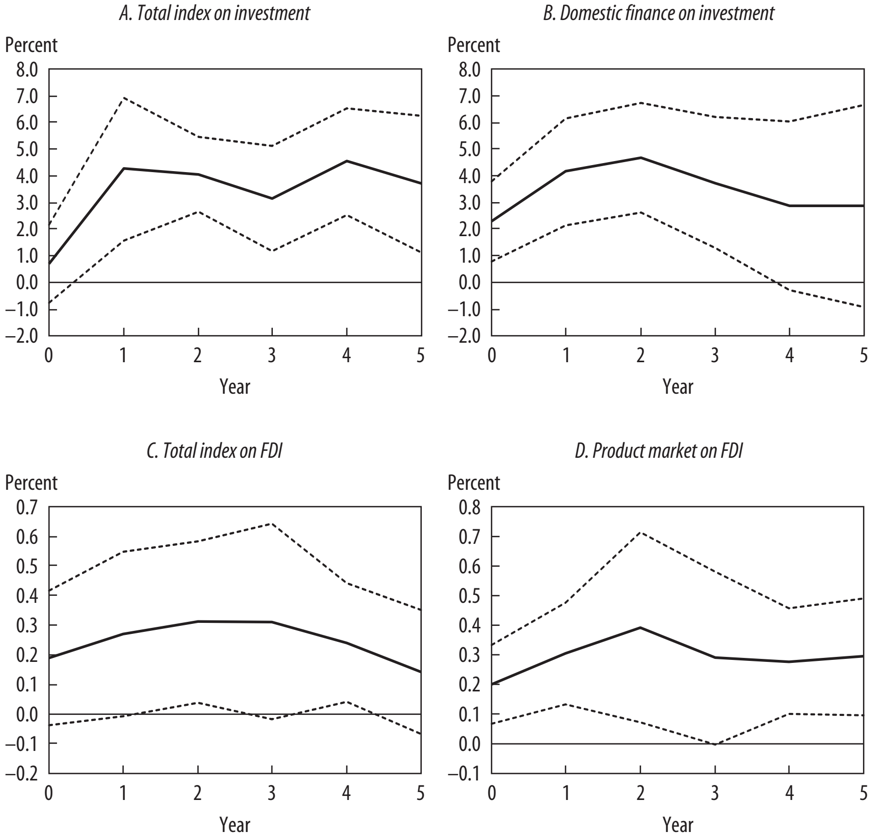


B. Product market and the labor index



Source: Authors' calculations, based on IMF data.
Note: The dashed lines show the 90 percent confidence interval for Driscoll-Kraay standard errors for the low index.

FIGURE 10. Effects of Structural Reforms on Investment and FDI



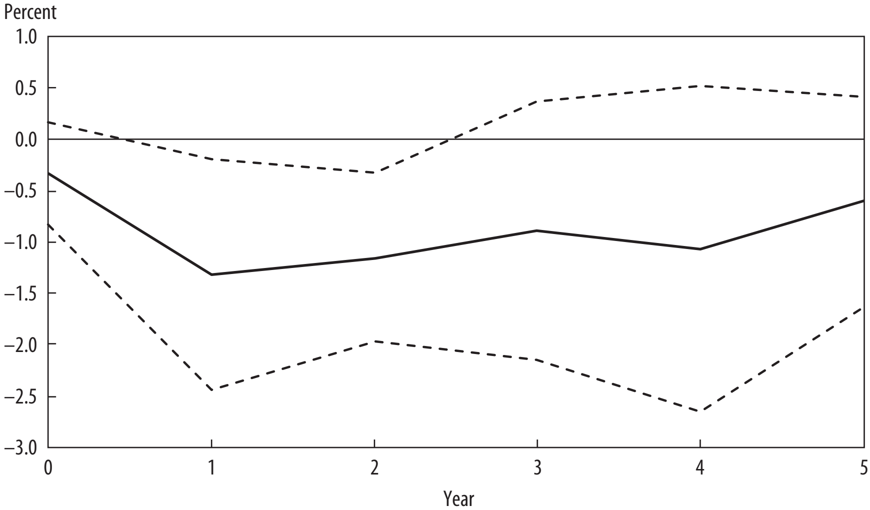
Source: Authors' calculations, based on IMF data.
Note: The dashed lines show the 90 percent confidence interval for Driscoll-Kraay standard errors.

rate by about one percentage point over five years. IMF (2019) also finds that large reforms lead to a reduction in informality rates of the same magnitude (about 1 percent over a five-year horizon) for a broader set of countries.

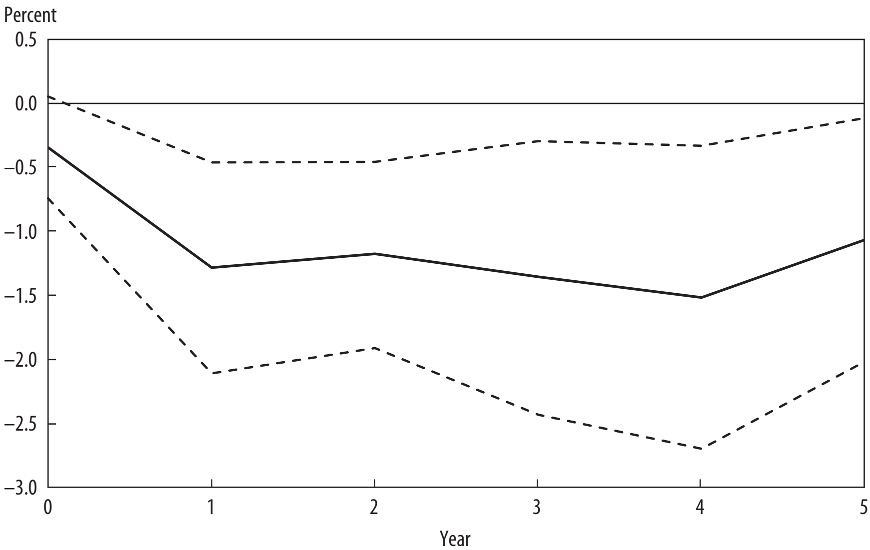
Policymakers frequently claim that reforms have important effects on business confidence, arguing that the boost in confidence associated with reforms could even offset the fiscal costs linked to their implementation. To tackle this issue, we estimate impulse responses for an index of business confidence from Haver Analytics for a sample of fourteen countries, including both advanced economies and emerging markets. We do not restrict ourselves to the sample

FIGURE 11. Effects of Structural Reforms on the Informality Rate

A. Total index



B. Product market index



Source: Authors' calculations, based on IMF data.
Note: The dashed lines show the 90 percent confidence interval for Driscoll-Kraay standard errors.

of Latin American and Caribbean economies in this case owing to the limited data availability for the business confidence indicators (the index is available in a comparable manner only for Brazil, Chile, Mexico, and Peru).

Figure 12 presents impulse responses for changes in the total reform index and in the employment protection index. Overall, the effects of large reforms on business confidence are positive, but not statistically significant. When focusing on reforms to job protection legislation, we find positive effects on confidence that take time to materialize, becoming apparent only two years after the changes in the reform index occur (there are forty-three changes in the labor reform index over the sample considered). Thus the data do not seem to support the view that reforms lead to substantial immediate improvements in business confidence. Effects can be positive and significant, but they seem to take time to materialize.

Structural Reforms and External Trade

We now turn to the effects of reforms on external trade. Overall, reforms boost growth in real exports over the medium term (figure 13), and, naturally, the effects of trade liberalization are particularly prominent, even if other reforms such as product market liberalization (not shown) also increase real exports. Reforms increase real imports by comparable magnitudes. These conclusions also hold when we consider the ratios of exports and imports to GDP rather than the real variables, suggesting that the growth accelerations of exports and imports following reforms are larger than the acceleration of GDP growth.

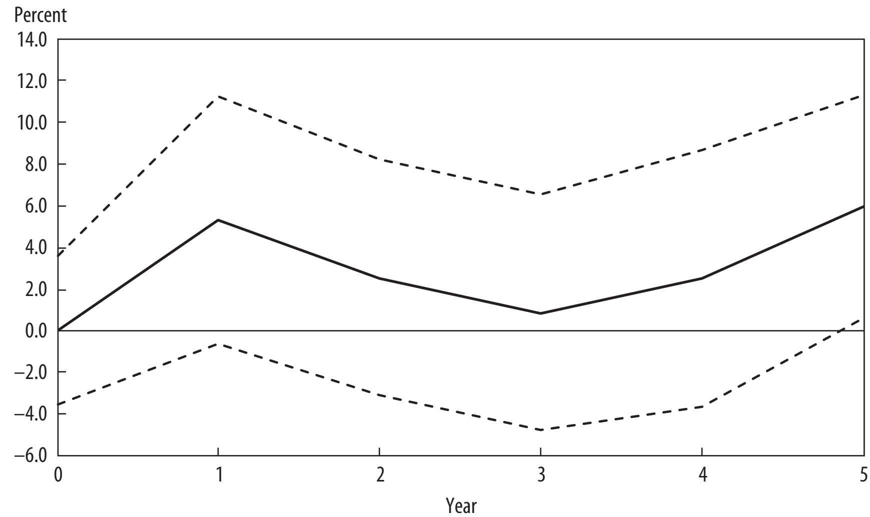
Reforms also appear to contribute to export diversification (figure 14). The Theil index for exports (a measure of concentration) declines after reforms, in particular after trade liberalization. This supports an argument frequently advanced in the international trade literature that high tariffs introduce an anti-export bias in some sectors, which liberalization appears to remove.

The Sectoral Effects of Reforms

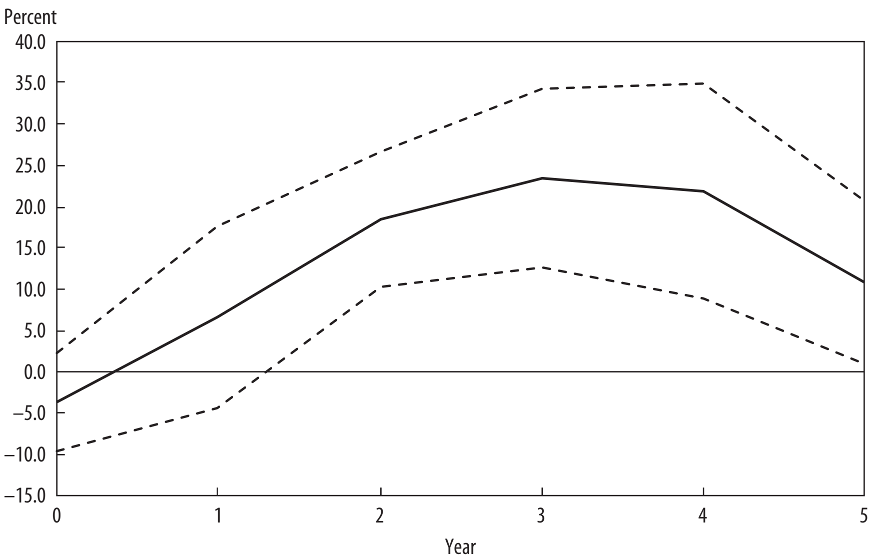
Liberalization could disproportionately affect specific sectors relative to others if the reforms relax important distortions or constraints on those sectors. The results show that changes in the aggregate reform index lead to increased real value added in manufacturing and agriculture (figure 15). In contrast, the effects of reforms on real value added in services are not statistically different from zero (figure 16). This suggests that reforms tend to favor tradable sectors.

FIGURE 12. Effects of Structural Reforms on Business Confidence: Fourteeen Economies

A. Total index

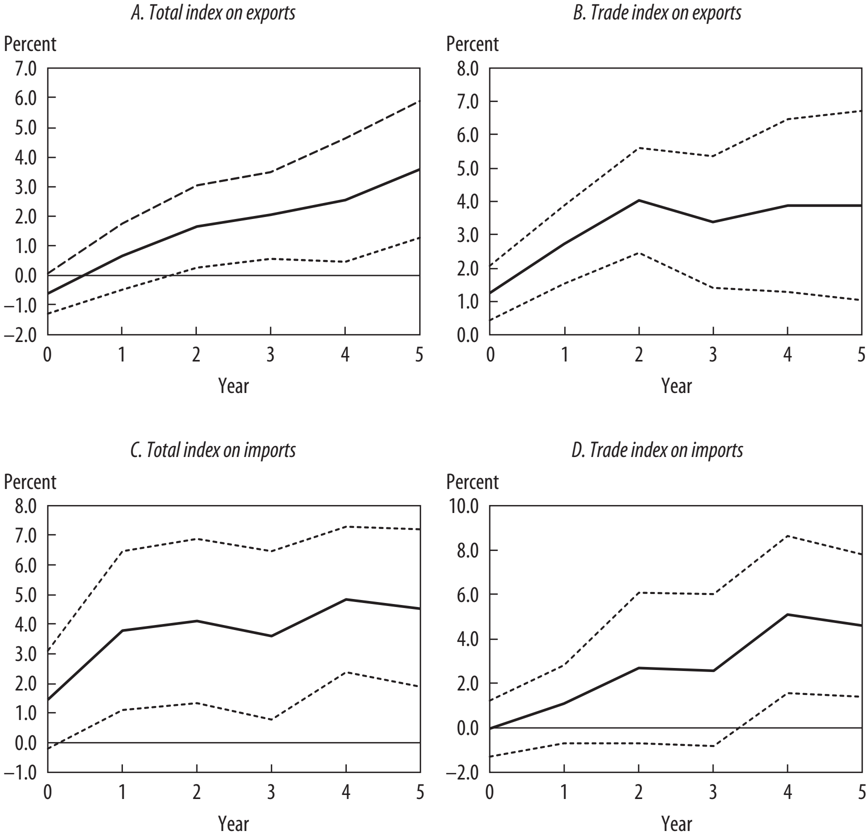


B. Labor index



Source: Authors' calculations, based on data from Haver Analytics.
Note: The dashed lines show the 90 percent confidence interval for Driscoll-Kraay standard errors.

FIGURE 13. Effects of Structural Reforms on Exports and Imports



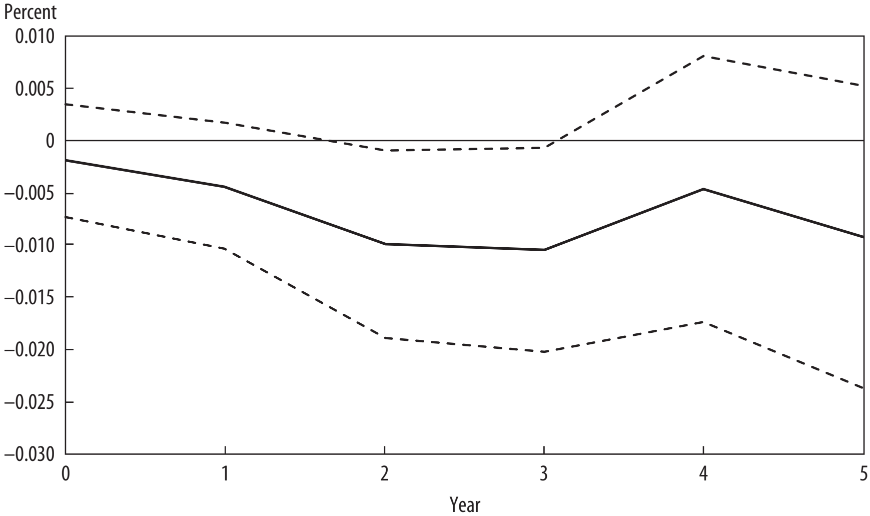
Source: Authors' calculations, based on IMF data.
Note: The dashed lines show the 90 percent confidence interval for Driscoll-Kraay standard errors.

As with GDP, the effects on manufacturing and agriculture value added tend to be significantly different from zero about two to three years after the reforms are implemented.

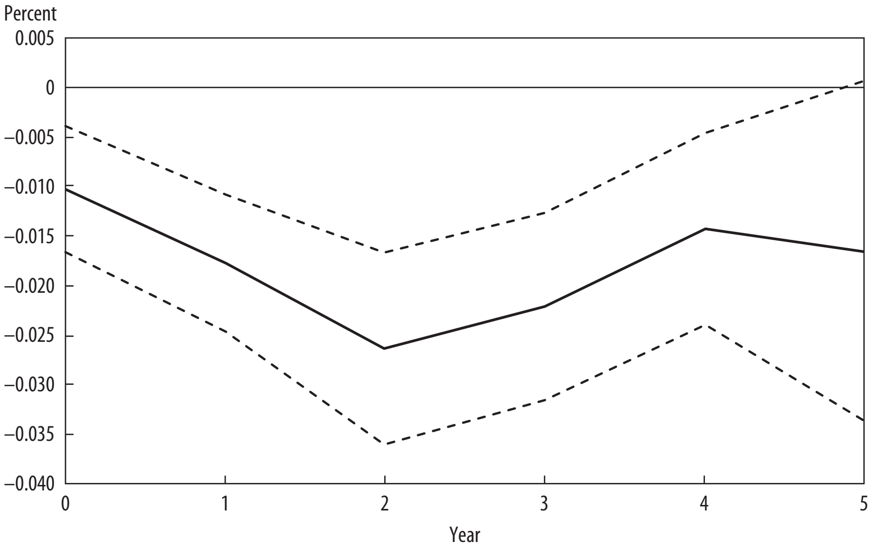
When we consider specific reform subindexes, it appears that each sector is affected by different reform clusters. Manufacturing value added increases after trade and product market reforms, while agriculture and services value added tend to increase following domestic finance and trade liberalization (figure 17).

FIGURE 14. Effects of Structural Reforms on Export Concentration: Theil Index

A. Total index



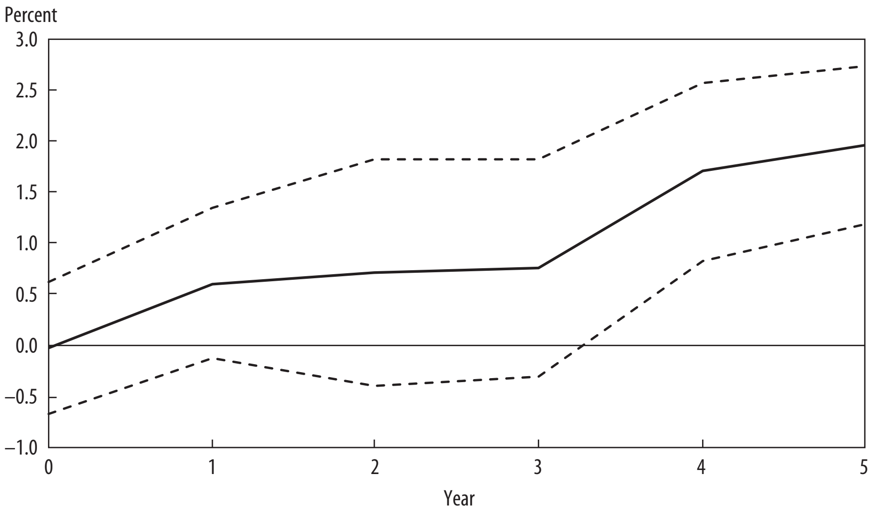
B. Trade index



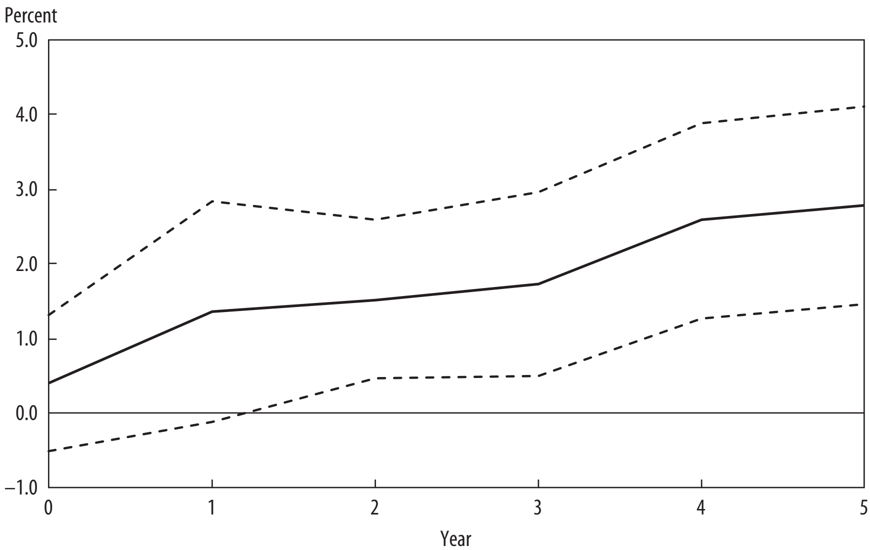
Source: Authors' calculations, based on IMF data.
Note: The dashed lines show the 90 percent confidence interval for Driscoll-Kraay standard errors.

FIGURE 15. Effects of Structural Reforms on Agriculture and Manufacturing Value Added

A. Total index on agriculture

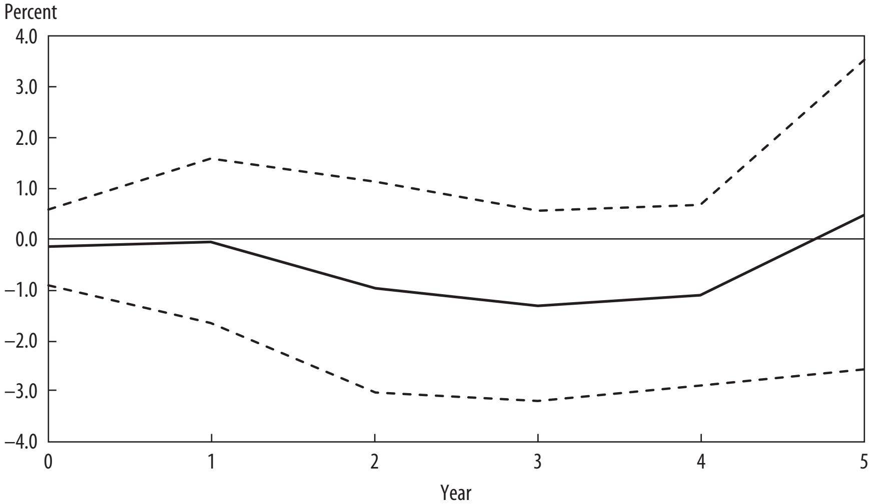


B. Total index on manufacturing



Source: Authors' calculations, based on IMF data.
Note: The dashed lines show the 90 percent confidence interval for Driscoll-Kraay standard errors.

FIGURE 16. Effects of Aggregate Structural Reforms on Services Value Added



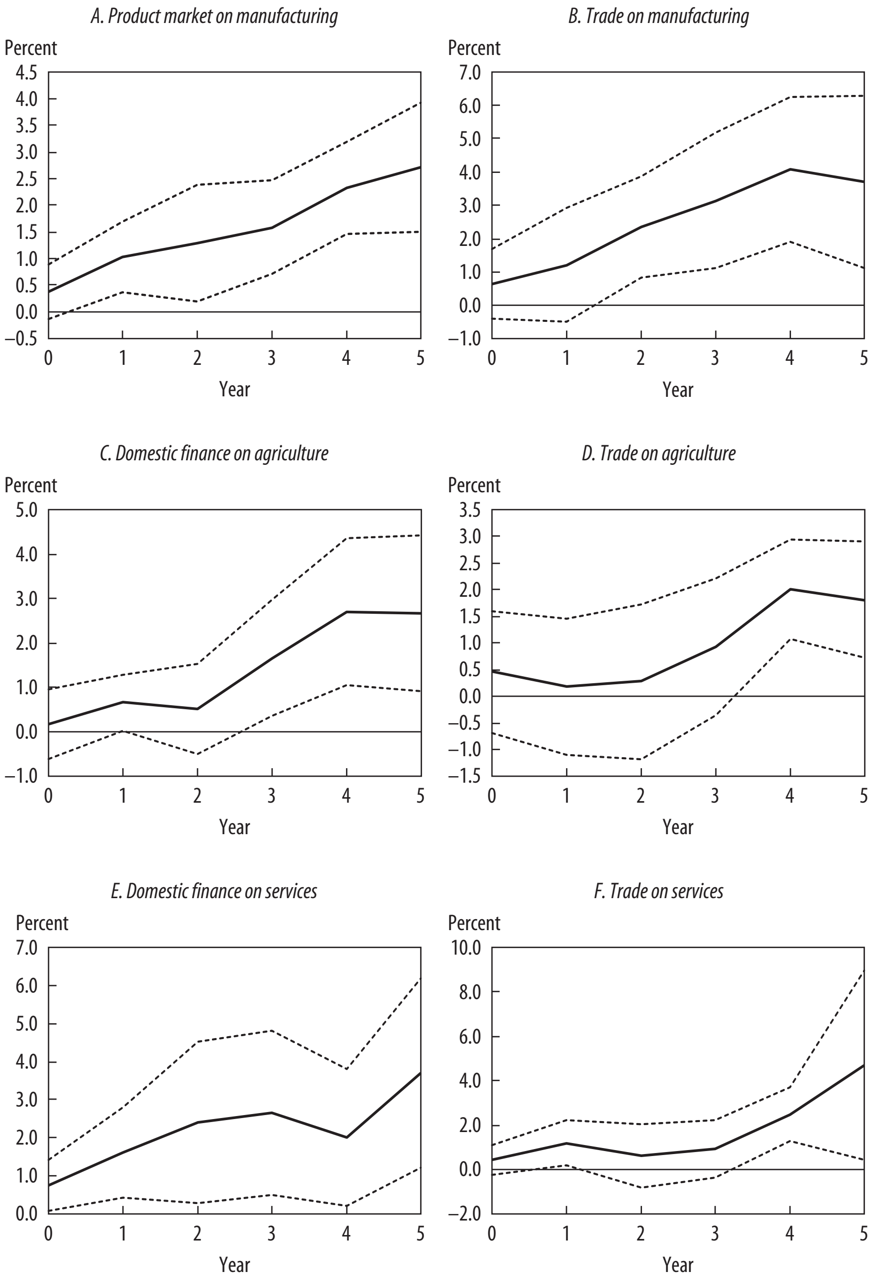
Source: Authors' calculations, based on IMF data.
Note: The dashed lines show the 90 percent confidence interval for Driscoll-Kraay standard errors.

Collateral Damage? The Effects of Reforms on Poverty and Inequality

Reforms are likely to affect different segments of society in distinct ways, which may partly explain resistance to reforms and subsequent reversals. This section examines whether market-oriented reforms might have deleterious effects over the short to medium term on inequality and poverty indicators in our sample of Latin American and Caribbean countries, using the same econometric framework outlined in previous sections. To measure the effects on the poverty rate, we use data on the poverty headcount ratio at USD 3.20 a day (2011 purchasing power parity) from the World Development Indicators database. To assess the effects on inequality, we use the Gini index from the same source.

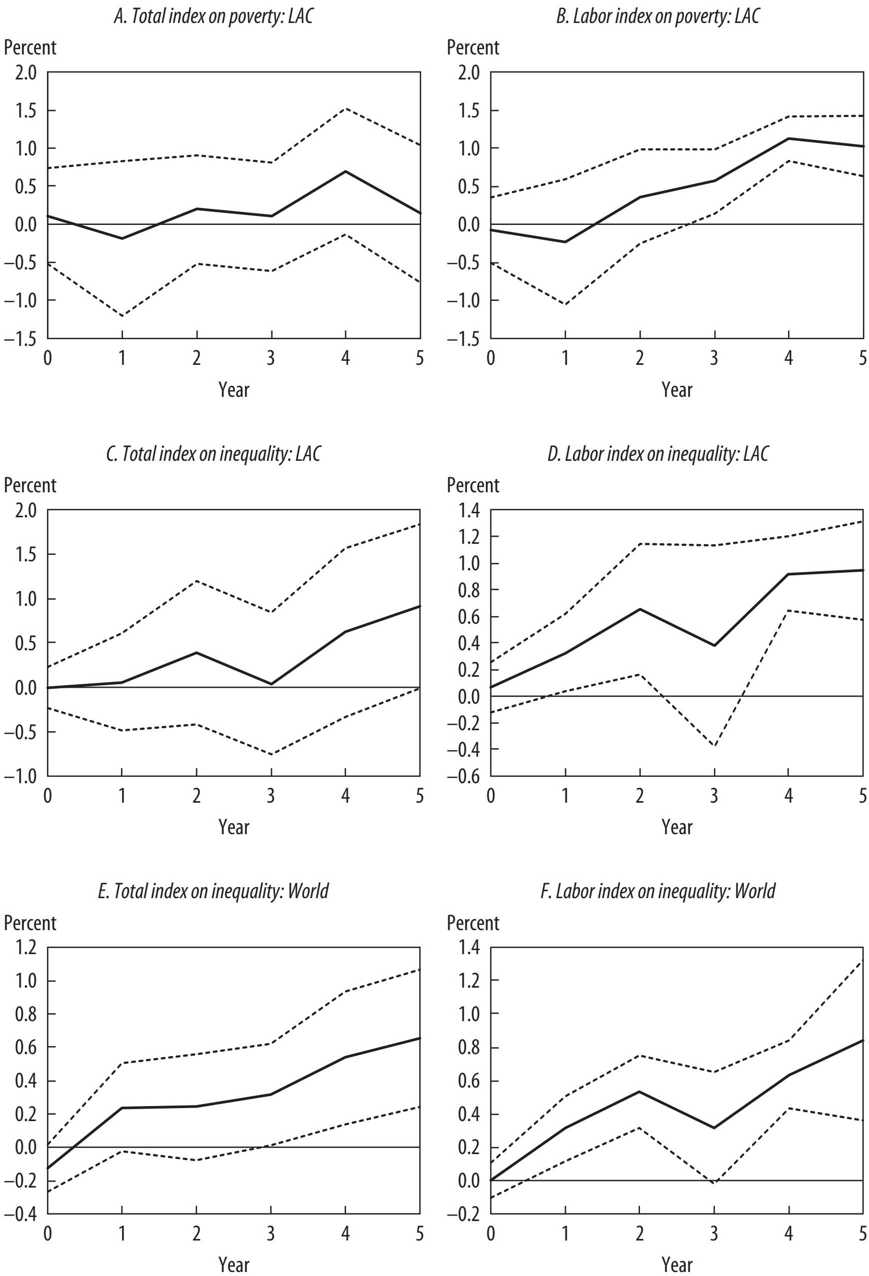
As illustrated in figure 18, we do not find statistically significant effects for the total reform index on poverty and inequality in our sample of countries. Nevertheless, reforms to job protection legislation are associated with statistically significant increases in both poverty and inequality indicators over the medium term. These effects appear to be economically small. Large changes in the employment protection reform index lead to increases in poverty rates of about one percentage point over five years. Similarly, inequality increases by

FIGURE 17. Sectoral Effects of Structural Reforms



Source: Authors' calculations, based on IMF data.
Note: The dashed lines show the 90 percent confidence interval for Driscoll-Kraay standard errors.

FIGURE 18. Effects of Structural Reforms on Poverty and Inequality



Source: Authors' calculations, based on World Development Indicators data.
Note: The dashed lines show the 90 percent confidence interval for Driscoll-Kraay standard errors. LAC: Latin America and the Caribbean.

about 1 percent over the same period. The last two panels of the figure reproduce the inequality regressions for the full sample of countries and confirm some of the results obtained for Latin America and the Caribbean. In the case of total reforms, the deleterious effects on inequality are now statistically significant, but remain of similar magnitude to the ones reported for Latin America and the Caribbean over the medium term.

Furceri and Rehman (2020) also report that reforms can be linked to increases in the Gini index when reforming countries have low intergenerational mobility and uneven access to opportunities, although the coefficients reported by these authors are smaller than our results for employment protection reforms. Nevertheless, they argue that for countries with high mobility and broad access to opportunities, the correlation between reforms and inequality tends to be insignificant or negative.

These results underscore the need to consider policy instruments to mitigate the potential negative effects of reforms, including measures to enhance access to opportunities. They strengthen the case for accompanying job protection liberalization with measures that protect workers, such as extending unemployment insurance schemes, as discussed in Duval and Loungani (2019).

Addressing Endogeneity Concerns

The results presented thus far have been interpreted as a causal relationship running from changes in the structural reform index to the specific variable of interest, but this interpretation is subject to caveats. One concern is that episodes of liberalization or reform reversals may be caused by past economic performance. For example, reforms may be implemented to revamp growth in countries experiencing a persistent economic slump. If this were the case, our results for the effect of liberalization on GDP could be picking up the persistence of growth. While this is partly captured in our baseline specification by including past growth as a control variable, there may be nonlinearities that we are unable to capture. Countries may also opt to liberalize key input markets in anticipation of higher growth (Buera, Monge-Naranjo, and Primiceri, 2011), which would also contaminate the causal interpretation of our results.

To address potential endogeneity concerns, we implement a panel instrumental variables (IV) strategy whereby we exploit the timing of liberalization and reform reversals across countries. More specifically, we instrument changes in the reform index in country i with current and past episodes of changes in the reform index in nearby countries. This identification strategy (namely, identification through regional waves) has also been used to study the causal effects of democratization on growth (Acemoglu and others, 2019) and the impact of fiscal austerity on social unrest (Ponticelli and Voth, 2020). In the specific case of episodes of liberalization and reform reversals, the exercise is grounded in the theoretical findings of Buera, Monge-Naranjo, and Primiceri (2011), who show that the adoption of liberalization policies by neighboring countries affects the belief of policymakers about the desirability of reforms.

The validity of the strategy rests on two assumptions: that regional waves are not affected by regional economic conditions (rather, they reflect regional demand for reforms that is unrelated to economic conditions) and that regional reform waves affect economic performance only through their impact on a country's adoption of reforms. While carefully assessing the validity of these assumptions goes beyond the scope of this paper, there is some evidence that regional waves do not respond exclusively to regional economic conditions. For example, Bonhomme and Maresa (2015) find that transitions to democracy are correlated at the regional level, even after controlling for GDP. In the case of structural reforms, Birdsall, de la Torre, and Valencia Caicedo (2010) argue that the wave seen in Latin America in the 1990s reflected the international view that "economic prosperity could only be obtained by harnessing the power of markets" (p. 7).′

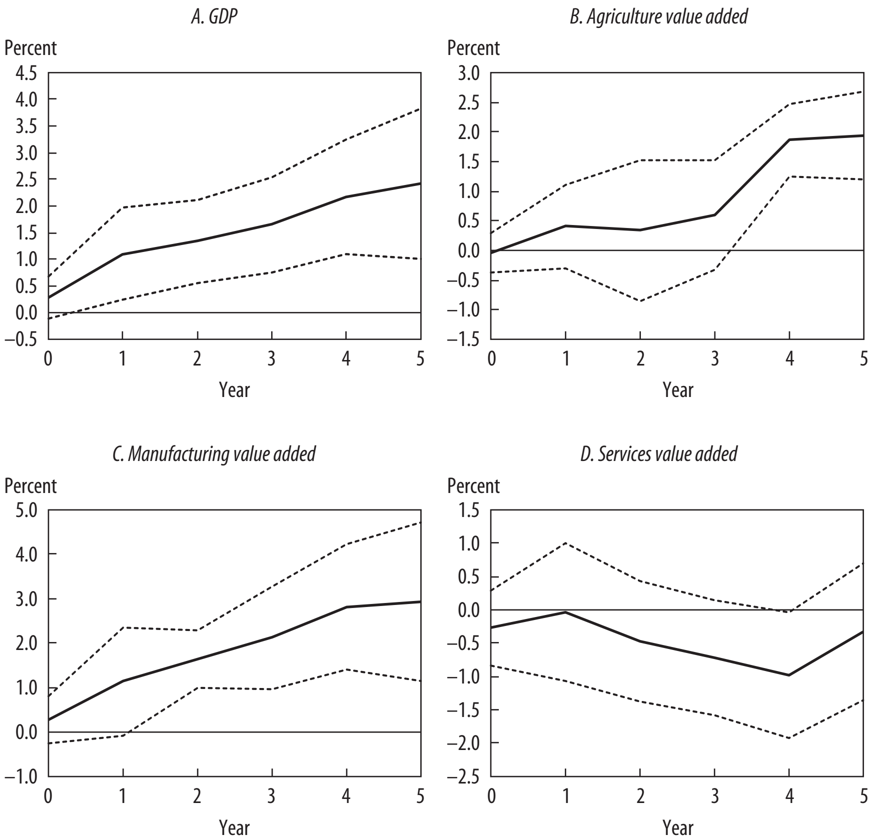
With this in mind, the empirical exercise presented in this section constructs a variable of changes in the reform index in nearby countries, as follows:

$$(3) \qquad \Delta SR_{i,t}^{-i,W} = \sum\nolimits_{j \in W_{-i}} \frac{1/\log\left(Dist_{i,j}\right)}{\sum\nolimits_{k \in W_{-i}} \left[1/\log\left(Dist_{i,k}\right)\right]} \Delta SR_{j,t},$$

where W_{-i} is the set of all countries for which we have data on the reform index excluding country i, $\Delta SR_{i,t}$ is the change in structural reform index, and $Dist_{i,j}$ is the bilateral population-weighted distance between country i and country j, as presented in the CEPII GeoDist data set. Once we construct the variable, we follow an IV strategy where $\Delta SR_{i,t}$ is instrumented using $\Delta SR_{i,t}^{-i,W}$ and its lagged values.[5]

5. In the IV exercise, we include two lags.

FIGURE 19. Effects of Structural Reforms: IV Approach

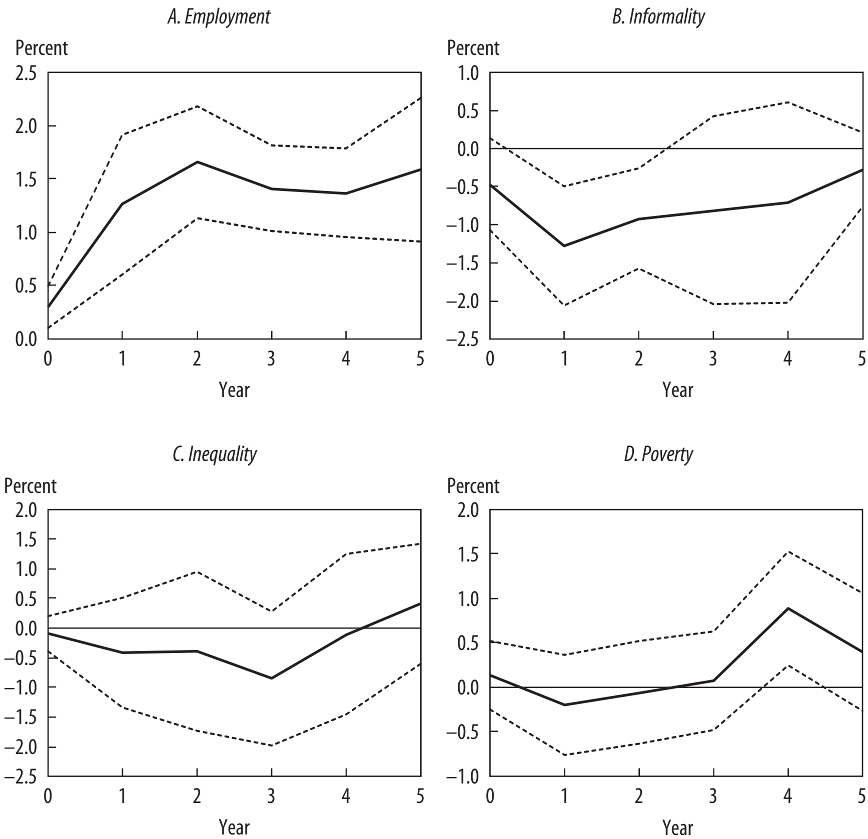


Source: Authors' calculations, based on IMF data.
Note: The dashed lines show the 90 percent confidence interval for Driscoll-Kraay standard errors.

The results, shown in figures 19 to 22, provide partial support to the causal interpretation of our findings.[6] Changes in the structural reform index associated with similar changes in nearby countries result in a gradual and statistically significant increase in GDP, employment, and trade outcomes. Informal employment decreases temporarily, and poverty appears to rise toward the end of the

6. The hypothesis of weak instruments is rejected. The Cragg-Donald F statistic for the first stage is 517.7, while the Kleibergen-Paap F statistic is 43.4. Both cases exceed the Stock-Yogo critical values.

FIGURE 20. Effects of Structural Reforms on Employment and Social Indicators: IV Approach

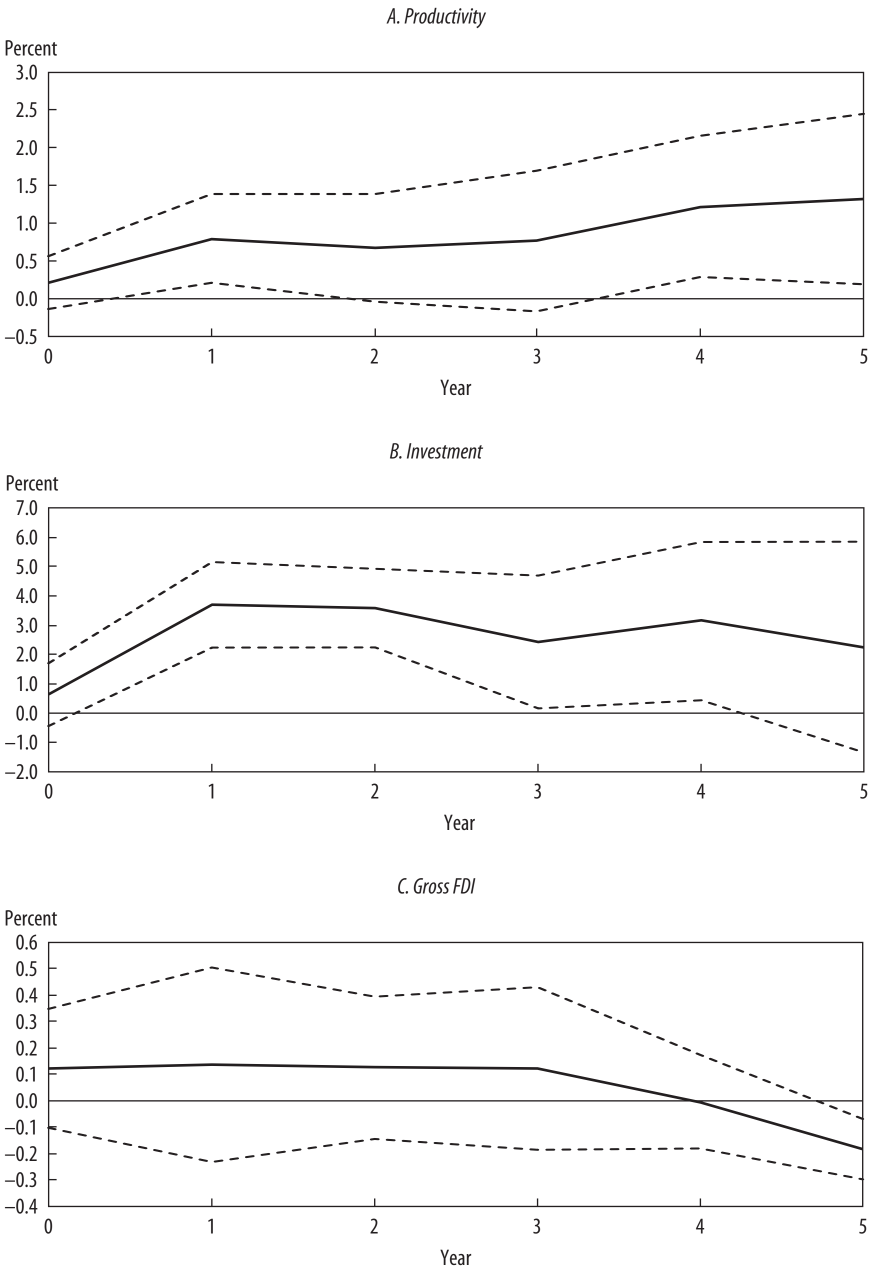


Source: Authors' calculations, based on IMF data.
Note: The dashed lines show the 90 percent confidence interval for Driscoll-Kraay standard errors.

window of analysis. As in the baseline results, the increase in GDP following episodes of liberalization appears to be driven by a surge in investment (demand dimension) and an increase in agricultural and manufacturing value added (sectoral dimension).

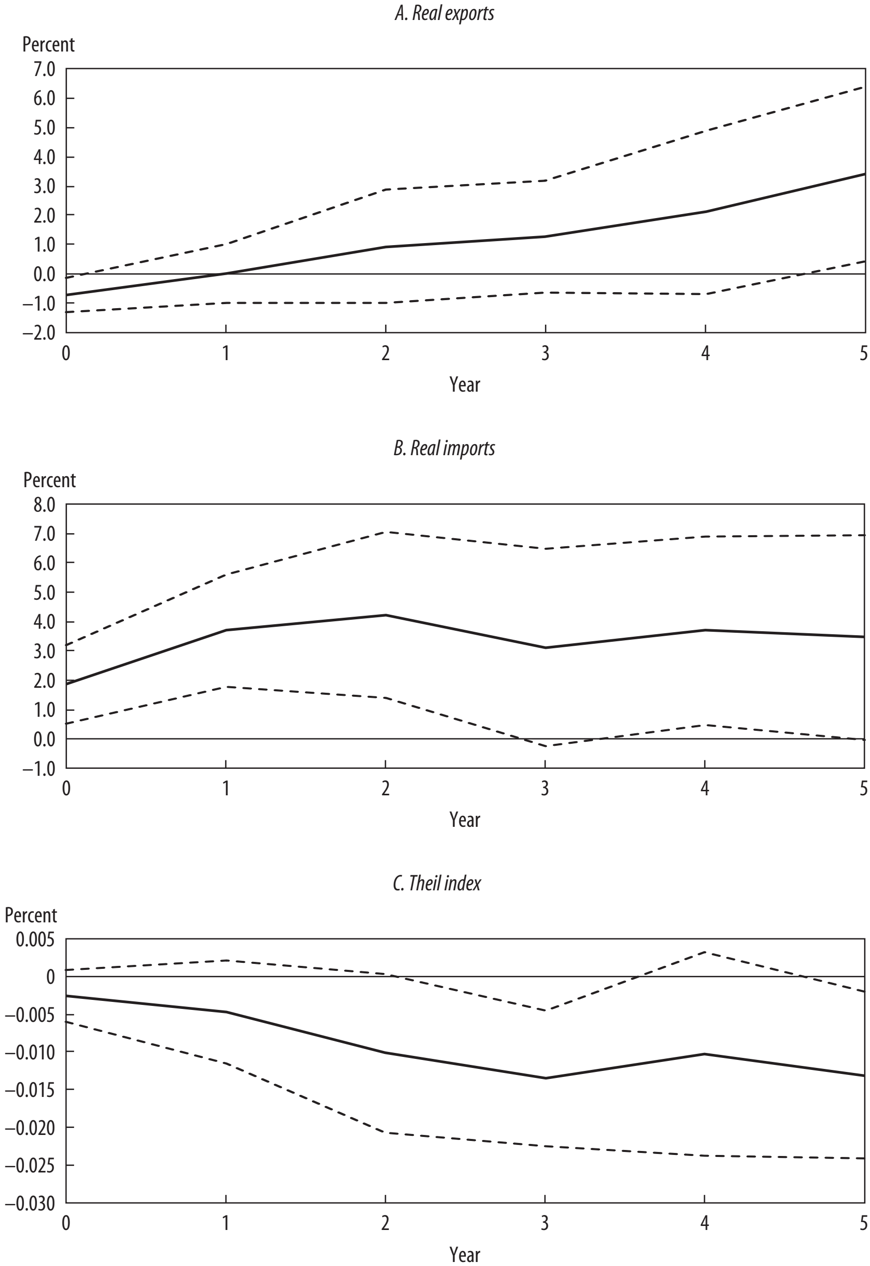
The IV estimates are similar in magnitude to the baseline exercise, but the significance of the results varies depending on the outcome of interest. For example, the IV approach shows that productivity increases following an increase in the structural reform index, but this increase takes four years to materialize (in the sense of statistical significance). The baseline exercise

FIGURE 21. Effects of Structural Reforms on Productivity and Investment: IV Approach



Source: Authors' calculations, based on IMF data.
Note: The dashed lines show the 90 percent confidence interval for Driscoll-Kraay standard errors.

FIGURE 22. Effects of Structural Reforms on Trade Outcomes: IV Approach



Source: Authors' calculations, based on IMF data.
Note: The dashed lines show the 90 percent confidence interval for Driscoll-Kraay standard errors.

shows that TFP increases after liberalization, but the effect is (marginally) statistically significant after five years. For real exports, we get the opposite pattern: the impact of changes in the structural reform index in the IV exercise is significant only five years after the change, while in the baseline exercise we get statistical significance after two years.

Conclusion

Throughout this paper, we have provided evidence suggesting that structural reforms have had broadly positive macroeconomic effects on a number of dimensions in countries in Latin America and the Caribbean. Nevertheless, reforms in some areas remain highly controversial in the region. Despite the potentially positive effects of trade, product market, and financial market reforms documented in this paper, there is still significant resistance from the public in the region toward reform efforts in these areas, in particular as far as trade liberalization is concerned. This could be explained by a number of factors considered in the political economy literature, including uncertainty regarding the winners and losers from reform (Alesina and Drazen, 1991; Fernández and Rodrik, 1991).

Another potential obstacle could stem from concerns about the reforms' effects on electoral outcomes and their associated political costs. Alesina and others (2020) find that timing matters in this regard. If reforms are implemented early in the political cycle, they are less likely to entail electoral costs. In fact, empirically, reforms are only associated with electoral costs if they are implemented the year before the election. Overall economic conditions also matter. Reforms implemented in periods of strong economic activity typically are not penalized by the electorate.

The paper also discussed evidence that the effects of reforms are not uniform across segments of the population. In that context, the adoption of policies to mitigate the adverse effects of reforms is crucial and will help to foster sustainability. For example, when discussing reforms to liberalize labor markets, Duval and Loungani (2019) highlight the importance of strengthening unemployment insurance and other social benefits at the same time to guarantee adequate protection of workers. The tax system could also be used to redistribute some of the gains from reform. In most countries in the region, there is scope to increase the role of progressive (nonlinear) personal income taxes in the tax structure, while enhancing redistributive policies on the expenditure side.

Appendix: Data Sources and Definitions

Business confidence: Log of index number. A value over 100 is optimistic. Source: Haver Analytics.

Commodity terms of trade: Log of commodity terms-of-trade index. Source: Gruss and Kebhaj (2019).

Employment: Log of employment (in thousands of people). Source: World Economic Outlook database.

Export concentration: Theil entropy index. Source: Authors' calculations, based on the Standard International Trade Classification (SITC), Revision 2.

Foreign direct investment: FDI inflows, percent of GDP in U.S. dollars. Source: IMF Financial Flows Analytics (FFA) database.

Inequality index: Gini index (World Bank estimate). Source: World Development Indicators.

Informality rate: Share of active workers not contributing to social security. Source: IDB Social Security Information System (SIMS) database.

Poverty rate: Poverty headcount ratio at USD 3.20 a day (2011 purchasing power parity). Source: World Development Indicators.

Real exports and imports: Exports and imports of goods and services, in 2010 U.S. dollars. Source: World Development Indicators.

Real GDP: Log of real GDP (in billions of local currency units). Source: World Economic Outlook database.

Real investment: Log of real gross capital formation (in billions of local currency units). Source: World Economic Outlook database.

Real sectoral value added: Sectoral value added, in constant 2010 U.S. dollars. Source: World Development Indicators

Structural reform index:. See main text. Source: Alesina and others (2020) and authors' calculations.

Total factor productivity: TFP at constant national prices (2011 = 1). Source: Feenstra, Inklaar, and Timmer (2015).

References

Alesina, Alberto F., and Allan Drazen. 1991. “Why Are Stabilizations Delayed?” *American Economic Review* 81 (5): 1170–88.

Alesina, Alberto F., Davide Furceri, Jonathan D. Ostry, and others. 2020. “Structural Reforms and Electoral Outcomes: Evidence from a World-Wide New Dataset.” Working Paper 26720. Cambridge, MA: National Bureau of Economic Research.

Acemoglu, Daron, Suresh Naidu, Pascual Restrepo, and James A. Robinson. 2019. “Democracy Does Cause Growth.” *Journal of Political Economy* 127 (1): 47–100.

Bakker, Bas B., Manuk Ghazanchyan, Alex Ho, and Vibha Nanda. 2020. “The Lack of Convergence of Latin America Compared with CESEE: Is Low Investment to Blame?” Working Paper 20/98. Washington, DC: IMF.

Birdsall, Nancy, Augusto de la Torre, and Felipe Valencia Caicedo. 2010. “The Washington Consensus: Assessing a Damaged Brand.” Policy Research Paper 5316. Washington, DC: World Bank.

Bergoeing, Raphael, Patrick J. Kehoe, Timothy J. Kehoe, and Raimundo Soto. 2001. “A Decade Lost and Found: Mexico and Chile in the 1980s.” Working Paper 8520. Cambridge, MA: National Bureau of Economic Research.

Biljanovska, Nina, and Damiano Sandri. 2018. “Structural Reform Priorities for Brazil.” Working Paper 18/224. Washington, DC: IMF.

Billmeier, Andreas, and Tommaso Nannicini. 2013. “Assessing Economic Liberalization Episodes: A Synthetic Control Approach.” *Review of Economics and Statistics* 95 (3): 983–1001.

Bonhomme, Stéphane, and Elena Maresa. 2015. “Grouped Patterns of Heterogeneity in Panel Data.” *Econometrica* 83 (3): 1147–84.

Buera, Francisco J., Alexander Monge-Naranjo, and Giorgio Primiceri. 2011. “Learning the Wealth of Nations.” *Econometrica* 79 (1): 1–45.

Ciminelli, Gabriele, Davide Furceri, Jun Ge, and others. 2019. “The Political Costs of Reforms: Fear or Reality?” Staff Discussion Note 19/08. Washington, DC: IMF.

David, Antonio C., Jorge Roldos, and Samuel Pienknagura. 2020. “Labor Market Dynamics, Informality, and Regulations in Latin America.” Working Paper 19/230. Washington, DC: IMF.

Dias Da Silva, António, Audrey Givone, and David Sondermann. 2017. “When Do Countries Implement Structural Reforms?” ECB Working Paper 2078. Frankfurt: European Central Bank.

Duval, Romain A., Davide Furceri, and Jakob Miethe, J. 2021. “Robust Political Economy Correlates of Major Product and Labor Market Reforms in Advanced Economies: Evidence from BAMLE for Logit Models.” *Journal of Applied Econometrics* 36 (1): 98–124.

Duval, Romain A., and Prakash Loungani. 2019. “Designing Labor Market Institutions in Emerging and Developing Economies: Evidence and Policy Options.” Staff Discussion Note 19/04. Washington, DC: IMF.

Feenstra, Robert C., Robert Inklaar, and Marcel P. Timmer. 2015. "The Next Generation of the Penn World Table." *American Economic Review* 105 (10): 3150–82.

Fernández, Andrés, Andrés González, and Diego Rodríguez. 2018. "Sharing a Ride on the Commodities Roller Coaster: Common Factors in Business Cycles of Emerging Economies." *Journal of International Economics* 111 (March): 99–121.

Fernández, Raquel, and Dani Rodrik. 1991. "Resistance to Reform: Status Quo Bias in the Presence of Individual-Specific Uncertainty." *American Economic Review* 81(5): 1146–55.

Furceri, Davide, and Sidra Rehman. 2020. "What's Opportunity Got to Do with It?" Washington, DC: IMF.

Gruss, Bertrand, and Suhaib Kebhaj. 2019. "Commodity Terms of Trade: A New Database." Working Paper 19/21. Washington, DC: IMF.

IMF (International Monetary Fund). 2019. "Reigniting Growth in Emerging Market and Low-Income Economies: What Role for Structural Reforms?" In *World Economic Outlook, October 2019: Global Manufacturing Downturn, Rising Trade Barriers*, chap. 3. Washington, DC: IMF.

Jordà, Òscar, and Alan M. Taylor. 2016. "The Time for Austerity: Estimating the Average Treatment Effect of Fiscal Policy." *Economic Journal* 126 (590): 219–55.

Lora, Eduardo. 2012. "Las reformas estructurales en América Latina: Qué se ha reformado y cómo medirlo (versión actualizada)." Working Paper No. 346. Washington, DC: Inter-American Development Bank.

Ponticelli, Jacopo, and Hans-Joachim Voth. 2020. "Austerity and Anarchy: Budget Cuts and Social Unrest in Europe, 1919–2008." *Journal of Comparative Economics* 48 (1): 1–19.

Prati, Alessandro, Massimiliano G. Onorato, Chris Papageorgiou. 2013. "Which Reforms Work and under What Institutional Environment? Evidence from a New Data Set on Structural Reforms." *Review of Economics and Statistics* 95 (3): 946–68.

Zettelmeyer, Jeromin. 2006. "Growth and Reforms in Latin America: A Survey of Facts and Arguments." Working Paper 06/210. Washington, DC: IMF.

CHRISTOPHER O'LEARY
W. E. Upjohn Institute for Employment Research, United States

TÚLIO CRAVO
African Development Bank, Ivory Coast

ANA CRISTINA SIERRA
World Bank, United States

LEANDRO JUSTINO
World Bank, United States

Effects of Job Referrals on Labor Market Outcomes in Brazil

ABSTRACT This paper is the first to use program administrative data from Brazil's National Employment System (SINE) to assess the impact of SINE job interview referrals on labor market outcomes. We use data from a five-year period (2012–16) to evaluate the impact of SINE job referrals on reemployment, time until reemployment, job tenure, and wage rates. Causal impact estimates based on propensity score matching suggest that a SINE job interview referral increases the probability of finding a job within three months of the referral and reduces the number of months needed to find reemployment, the average job tenure of the next job, and the reemployment wage. Subgroup analysis suggests that SINE is particularly effective at helping less educated workers find work in a timely fashion. Finally, the evidence suggests that the self-service online labor exchange is less effective than the in-person job interview referrals provided at SINE offices.

JEL Codes: J18, J23, J68

Keywords: Labor market policy, employment services, job interview referrals, difference-in-differences

Countries in Latin America and the Caribbean faced an array of labor market problems in the 1990s, including high unemployment, poor working conditions, and a lack of quality job opportunities. This situation generated policy interest in improving labor market programs, especially the public labor exchange. In recent years, as labor market policy has become an important macroeconomic policy instrument in the region, labor market programs have garnered a bigger share of public resources and have served more job seekers and employers (Ramos, 2002).

In Brazil, labor markets have performed reasonably well over the past fifteen years in terms of labor market participation and labor earnings growth.

ACKNOWLEDGMENTS We are grateful to the reviewers and the editor for helpful comments. Financial support from the Inter-American Development Bank for an earlier version of the paper and from the W. E. Upjohn Institute for Employment Research is acknowledged. We thank the Ministry of Labor, Brazil, for providing access to the data.

However, a recession that started in the second quarter of 2014 nearly doubled the unemployment rate, from an average of 6.9 percent in 2011–2014 to an average of 12.0 percent in the subsequent four years.[1]

The country's National Employment System (SINE) is a key institution for public employment policies. Created in 1975, this network of local employment offices serves as a go-between, helping workers line up jobs and providing information to employers on available workers.[2] The Worker Protection Fund, established in 1990, expanded SINE to 1,930 offices in 2016, with locations throughout the country, covering all twenty-six states and the Federal District. The Ministry of Labor coordinates this large network, monitoring the decentralized delivery of services by states and municipalities.[3]

SINE customers tend to be less educated and lower skilled, but SINE also provides services for customers with higher educational attainment and job qualifications. In this paper, we estimate the program's causal impacts on the full range of customers and analyze the effects of job referrals on all customers, most of whom have work histories characterized by high rates of turnover in formal sector jobs. Our estimates, using propensity score matching (PSM) to create comparison groups and difference-in-differences estimators to compute impacts, suggest that SINE job referrals increase the probability of finding a job and reduce the time to reemployment, the average tenure in the next job, and the reemployment wage. Our subgroup analysis further suggests that SINE could broaden its impact by expanding services to more highly educated job seekers. We find that it takes almost twice as long (nine weeks) to fill a skilled job vacancy in Brazil as it does on average (five weeks) in other Latin American and Caribbean countries (Aedo and Walker, 2012).

Improving the effectiveness of the public employment service (PES) is essential to supporting quick, successful, and durable job matches (Betcherman, Olivas, and Dar, 2004). An effective PES contributes to labor market efficiency, reducing informational breakdowns that slow or prevent the proper matching of job seekers' skills to employers' job vacancies. Borges, Lobo, and Foguel

1. According to the Brazilian Business Cycle Dating Committee (CODACE) of the Brazilian Institute of Economics (IBRE), the recession lasted eleven quarters, from the second quarter of 2014 to the last quarter of 2016.

2. SINE was created after the Brazilian government ratified Convention No. 88 of the International Labour Organization (ILO), which relates to the organization of public employment services. SINE is also one of the means through which workers request unemployment benefits. For more details on SINE, see IPEA (2020) and Lobo and Anze (2016).

3. The Ministry of Labor was integrated into the Ministry of Economy following the restructuring of the federal ministries in 2019. The Secretariat of Productivity, Employment, and Competitiveness in the Ministry of Labor is currently responsible for the SINE network.

(2017) estimate that PES labor intermediation in Brazil saved the Worker Protection Fund about R$43 million in 2016 through reduced unemployment insurance payments. Since labor intermediation programs typically benefit low-skilled workers, countries with a large proportion of such job seekers could benefit from increased investment in labor exchange services.

As a percentage of the total budget for all active labor market programs, spending on labor intermediation services in Brazil is low compared to members of the Organization for Economic Cooperation and Development (OECD): Brazil spends less than 2 percent on labor intermediation services, while OECD countries spend an average of 10 percent (Silva, Almeida, and Strokova, 2015). Since the PES provides services free of charge, it also improves equity in access to social participation through the labor market. Although not an explicitly stated organizational objective, the movement of workers from informal to formal sector jobs might provide access to private health insurance and other benefits. Even if labor intermediation does not have a significant effect on aggregate employment, it can help maintain the attachment of the long-term unemployed to the labor force, thereby decreasing their dependence on social assistance programs.

When one considers the importance of public employment services, the paucity of research on program effectiveness in developing countries is remarkable. The studies conducted in the United States and Europe consistently find positive evidence of effectiveness for public labor exchange services in those developed countries (Blundell and others, 2004; Johnson, Dickinson, and West, 1985; Michaelides and Mueser, 2018). While the estimated impacts on employment and earnings are typically small, the low cost of interventions often makes PES job search assistance services cost-effective.

The few studies from Latin America showing causal evidence from survey data provide mixed results. Vera (2013), based on a small survey of 150 job applicants, finds that PES participation in Peru lengthens unemployment spells by thirty-three days. Pignatti (2016), using a nationwide survey for Colombia, finds that the Colombian PES increased participants' likelihood of having a formal job by between five and thirty-one percentage points, but had a small negative effect on hourly earnings, which declined between 2 and 5 percent.

While high-quality statistics on the administration of nationwide programs for labor intermediation in Brazil exist, to date there has not been a formal impact evaluation. This paper is the first study in Latin America to use a large body of observational data to produce a more robust evaluation of a labor intermediation service. Using administrative microdata from 2012 to 2016, our study uses PSM to create comparison groups and difference-in-differences estimators to compute impacts of SINE job referrals on labor market outcomes. Our difference-in-differences estimates suggest that a job referral by SINE increases employment probability within the next three months and reduces the number of months until employment. However, we also find that SINE referrals decrease the average tenure and wage of the next job. Our paper shows two other things: SINE job referral impacts differ across subgroups, and web-based job interview referrals contribute to the placement of workers but are less effective than face-to-face services in shortening nonemployment spells. Knowledge of these results can help program administrators design strategies to improve labor intermediation services.

The remainder of this paper is structured as follows. After summarizing the related literature, we describe the data used in the analysis and present summary statistics. Subsequent sections detail our methodology and present our results. The final section offers concluding remarks.

Background

Previous researchers provide mixed evidence on the effectiveness of work intermediation programs. Evaluations of the PES have focused mainly on the service's impacts on employment probability, unemployment duration, and earnings. Attempts to assess the impact of job interview referrals in the United States and Europe date back to the 1980s (for example, Johnson, Dickinson, and West, 1985; Jacobson and Petta, 2000), but the early U.S. studies did not provide convincing causal evidence of effectiveness.

More recently, Blundell and others (2004) used differences in the geographic location and demographic targeting of services to convincingly identify the effect of the New Deal for Young People program in the United Kingdom, which provided compulsory job search assistance to unemployment compensation applicants and wage subsidies to employers. The authors provide causal evidence that job search assistance increased the probability of young men finding a job in the next four months by five percentage points. This impact diminished over time, perhaps because of displacement effects.

Crépon and others (2013) used randomized controlled trials in a field experiment to measure the impacts of job placement assistance on the labor market outcomes of young, educated job seekers in France. They provide strong causal evidence that even though the program increased the likelihood of finding a stable job, the positive effect diminished over time and often came at the expense of other eligible workers. However, the SINE facilitates only about 3 percent of job placements, suggesting that displacement effects are a smaller concern in Brazil.

A more recent randomized controlled trial in the United States during the Great Recession identified unemployment insurance applicants who were likely to exhaust benefits and randomly assigned them to eligibility assessment, job search assistance, or nothing (Michaelides and Mueser, 2018). Strong causal evidence suggests that the treatment group had a 15 percent lower rate of exhausting regular unemployment benefits and an average 7.0 and 8.2 percentage point higher reemployment rate one and two quarters after treatment assignment, respectively. The results suggest that actions targeting unemployment insurance recipients can enhance labor intermediation services.

Few studies explore the effectiveness of PES agencies in South America. Vera (2013) conducted one study in Peru using a quasi-experimental design and finds that job search assistance provided by the Peruvian PES had only small impacts on unemployment spells compared with job search assistance from private agencies. However, her research design has important limitations for generating convincing causal evidence: the treated sample was based on information on program beneficiaries collected from a survey distributed to only 150 job applicants whom the PES had placed in a job in September 2004.

Pignatti (2016) used PSM to identify causal effects of job placements by the Colombian PES relative to job placements by other means such as private agencies, public posting of job openings, newspaper or website advertisements, or family and friends. Based on data from the annual household survey (*Gran Encuesta Integrada de Hogares*) conducted by the National Administrative Department for Statistics, the study finds evidence suggesting that using the Colombian PES positively affects the probability of having a formal sector job. However, it also finds that PES job placements reduce earnings in Colombia. A limitation to the identification strategy is that Pignatti's (2016) data were based on a sample of PES users from a general household survey, meaning the data do not have a panel structure and do not provide detailed information on previous job search history.

Our paper relies on the full population of all PES users in Brazil, merged with longitudinal data on employment and earnings from the Annual Social Information Report (*Relação Anual de Informações Sociais*, RAIS). It is, to our knowledge, the most complete evaluation of labor intermediation conducted in Latin America. Therefore, unlike previous analyses for Latin America, we are able to directly investigate the effects of program participation on the probability of finding a job, since our unique data set allowed us to follow job seekers' labor history both before and after the SINE job interview referral.

Only the prior study by Woltermann (2002) attempts to assess the effectiveness of job interview referrals on different groups of participants in Brazil. The study finds that the only significant channels for transition into formal sector jobs are directly contacting the employer, using connections through family and friends, and responding to advertisements. The study is based on the monthly employment surveys (PME) collected by the Brazilian Institute for Geography and Statistics (IBGE) and does not include data from Brazilian employment services.

Thus, although the literature from Europe and the United States provides more credible results about labor intermediation programs, the existing literature in Latin America does not provide convincing impact evaluations of the effectiveness of such programs on employment probability, earnings, time until reemployment, and job tenure. This paper constitutes the most comprehensive attempt to date to understand the effectiveness of these nationwide labor market programs in the Latin American context, using administrative data from Brazil for the first time.

Data and Descriptive Statistics

We constructed a unique data set, merging administrative data from SINE with data from the RAIS to analyze the effectiveness of labor intermediation in Brazil. SINE was established in 1975 as a public agency for labor market programs, including the labor exchange. Its original purpose was to promote labor intermediation, but currently its services include professional orientation, referral to qualification and training programs, job interview referrals, job placement, labor market information, issuance of formal worker-identification credentials, and some components of the unemployment insurance program, including benefit payments.[4]

The intermediation process involves the registration of workers and employers, recording of the employment histories of job seekers, and listing of job vacancies. The process of SINE labor intermediation begins with job search registration at a SINE office or through the SINE website. Based on information in the SINE database, the labor exchange officer explores possible job matches between the profiles of registered job seekers and listings

4. See the following website for more details: portalfat.mte.gov.br/programas-e-acoes-2/sistema-nacional-de-emprego-sine/.

of available jobs. The SINE job-matching expert then presents job interview opportunities to the job seeker that match his or her skills and experience profile and proceeds to offer any suitable job interview referrals.[5] Since May 2014, the SINE job interview referral system also allows job seekers to make an online self-referral if the worker meets the minimum requirements listed by the employer in the job vacancy posting.[6] Thus the SINE labor intermediation process entails matching job seeker profiles with the requirements of vacancies, referring workers to interviews based on the matching results, and capturing referral outcomes, which we use in this evaluation.

The SINE intermediation service also involves the management of job vacancy listings from the moment they are received to the moment they are filled or expire. The SINE database, used for research purposes here for the first time, contains socioeconomic information on workers from their registration forms (age, gender, education, and employment status), as well as information on employers and records of available job vacancies and job interview referrals (status of the referral, employer feedback, and type of service offered). The SINE database includes the individual's unique identification number (*Cadastro de Pessoas Físicas,* CPF), which allows us to track job seekers during the period of analysis.

The SINE data are complemented by RAIS annual administrative data compiled by the Labor Ministry of Brazil. These data include employment and earnings information on all formal sector firms and employed workers in a given year.[7] All formally registered firms in Brazil report annual information on their employees. The RAIS includes detailed information about the employer, the employee, and the employment relationship (including wage, tenure, type of employment, hiring and separation dates, and reason for separation). Importantly, RAIS is an employer-employee matched data set that can be linked to the SINE data set using the CPF.

For this paper, the RAIS data were available from 2011 through 2016. The RAIS data set is structured so that each observation represents an employment relationship containing the dates of hiring and separation. We use these

5. A worker who is a beneficiary of the unemployment insurance benefit cannot refuse an interview referral without having an acceptable excuse (Federal Law No. 7,998 of 1990).

6. In 2016, online self-referrals accounted for 16 percent of the total number of referrals (see table 1). IPEA (2014) shows details of the flow chart of the SINE labor intermediation process.

7. Severance payments are based on RAIS records; thus employers and workers have a strong incentive to submit the annual RAIS declaration. The Ministry of Labor estimates that RAIS coverage represents about 97 percent of the formal sector.

data to construct a monthly panel with information on each individual's employment status for that month. Our aim is to analyze the exit from unemployment (nonformal employment) of workers with past experience in formal sector jobs.[8] The panel data allow us to observe workers with more than one job at the same time—that is, multiple jobholders. Since job loss for a multiple jobholder does not result in full unemployment, our sample excludes workers who at some point in our data period had multiple simultaneous formal sector jobs.[9]

Since most workers who seek assistance from SINE are unemployed (94 percent), we restrict the analysis to workers who were separated from their jobs at some point before a job interview referral. In the panel based on RAIS information, a period between jobs is a period of nonemployment in the formal sector. Using the separation and hiring dates in RAIS, we create a panel of individuals with formal sector employment histories and at least one month of nonemployment in the formal sector.[10]

Overall, the study addresses unemployed individuals who were never multiple jobholders in the period analyzed, but who had at least one job in the RAIS before a job interview referral. Naturally, sequential job holding is permitted in our sample, because a new job *after* the job interview referral is a positive outcome (for example, reemployment wages, tenure in the next job).[11] The unemployment (or nonformal employment) periods correspond to the periods for individuals who were hired at some point during the time span of the panel after being separated.[12] In these data, we observe about 95,000 job interview referrals each month. The average reemployment job tenure is less

8. Outcomes are measured using RAIS records that encompass only formal workers.

9. Simultaneous jobs are defined as two or more jobs with durations (start and end dates) overlapping in time. This guarantees the fulfillment of the assumption that the period following a dismissal is, in fact, a period with no formal employment.

10. RAIS data include formal sector workers. We refer to nonemployment in the formal sector as unemployment.

11. We observe that a person who gets a referral in 2012 has a 90 percent probability of finding a formal sector job within the next five years. This means that for outcomes that require the observation of a job after the referral, restricting the panel to workers with at least one unemployment spell and a registry of formal employment after having been referred for a job interview by SINE retains most of the observations in our panel. For the last year of data, about 43 percent of workers who got referrals in 2016 got a job in that same year.

12. The resulting panel includes 29 million workers with at least one unemployment spell and a total of 41 million unemployment spells, as some workers have more than one unemployment spell.

TABLE 1. Descriptive Statistics of SINE Labor Intermediation, 2012–16

Year	Workers registered	Vacancies	Referrals	Workers placed	Placement rate (%)	Online referrals
2012	8,231,696	3,072,010	5,937,727	730,489	12	0
2013	7,480,241	3,597,192	6,745,416	838,320	12	0
2014	6,232,876	2,715,616	5,834,709	686,295	12	152,444
2015	5,185,316	1,758,888	4,900,375	616,497	13	243,167
2016	4,587,164	1,151,366	3,783,357	402,365	11	211,906
Total	31,717,293	12,295,072	27,201,584	3,273,966	12	607,517

Source: Authors' calculations, based on data from the Brazilian Ministry of Labor.
Note: The placement rate is equal to the ratio of workers placed to referrals.

than two years, suggesting that the available five-year time span for the data is sufficient to measure reemployment job tenure.[13]

Combining the SINE and RAIS data sets allows us to trace the duration of formal sector employment, time until reemployment, and earnings on the new job for individuals who look for employment through SINE agencies compared with those who use other job search methods. Table 1 provides descriptive statistics on the labor intermediation activities of SINE between 2012 and 2016. We chose this period because a new data system was established in 2012 that improved data quality and reliability significantly, according to the Ministry of Labor. Table 1 shows that the total number of unique workers registered in the SINE system reached 31.4 million for the 2012–16 period.[14] While 70 percent of the vacancies available at SINE have at least one job interview referral, only 28 percent of the vacancies were filled through a SINE job referral.[15] The overall placement rate (workers placed by referral) of SINE is about 12 percent throughout the period of analysis. Online self-service referrals were permitted starting in 2014.

13. The average job tenure in these data is exactly 19.6 months. The average job tenure for the formal private sector in Brazil is about 3.5 years, according to DIEESE (2016).

14. Table 1 shows the number of new SINE registrants per year. For instance, in 2016, 4,587,164 workers who had never registered with SINE did so. Thus, 31.7 million is the number of unique workers registered.

15. In the SINE system, one "vacancy" posted by an employer might represent more than one position. For instance, a firm might submit one vacancy requiring ten employees. On average, 3.8 positions are offered for each SINE vacancy. This average increases to 5.4 positions per vacancy when taking into account only the vacancies with at least one position filled. The data on vacancies, referrals, and workers placed are flows in each year.

TABLE 2. Descriptive Statistics for Job Seekers Referred by SINE, 2015

	Observations	
Statistic	*Employed*	*Unemployed*
Percent of total observations	6	94
Age, sample mean	24.1	31.7
Race (%)		
Indigenous	0	0
White	38	42
Dark	11	12
Yellow	1	1
Brown	49	45
Education (%)		
Illiterate	0	0
Incomplete middle school	9	15
Middle school graduate	6	11
Incomplete high school	29	14
High school graduate	46	49
Incomplete college	7	7
College graduate	2	3
Specialization	0	0
Advanced degree/PhD	0	0
Gender (%)		
Male	48	58
Female	52	42

Source: Authors' calculations, based on data from the Ministry of Labor.

To evaluate the impact of labor intermediation, we construct a monthly database with PSMs of job seekers getting referrals to other unemployed workers not getting referrals. We used data on only the first referral each month per unemployed job seeker, even if that individual was referred more than once in a month.[16]

Table 2 shows that 94 percent of the referrals are made for unemployed job seekers, which is the group of workers analyzed in this study. The average age of the workers referred by SINE is higher for the unemployed than for the employed, and the difference between the two groups is around seven years. The mean age of all SINE referrals is about thirty years old. While

16. The placement rate (workers placed by referral) that considers one referral per month is higher (16 percent) because the number of workers placed remains the same, but the number of referrals is lower than listed in table 1 (see online appendix A, table A1). (Supplementary material for this paper is available online at http://economia.lacea.org/contents.htm.)

TABLE 3. Descriptive Statistics of SINE Labor Intermediation by State, 2012–16

State	Workers registered	Offices per state	Vacancies	Referrals per office	Placements per office	Placement rate (%)
Acre	80,247	11	8,832	2,008	395	19.7
Alagoas	393,550	43	137,497	4,316	1,984	46.0
Amapá	83,460	12	12,673	1,461	118	8.1
Amazonas	453,945	29	140,717	5,074	1,428	28.1
Bahia	1,859,443	149	563,919	9,216	1,962	21.3
Ceará	931,723	135	643,526	10,014	2,870	28.7
Distrito Federal	501,929	26	233,878	41,793	2,492	6.0
Espírito Santo	642,186	34	185,039	11,152	792	7.1
Goiás	1,150,209	90	419,242	11,468	1,005	8.8
Maranhão	552,293	47	49,209	1,990	674	33.8
Mato Grosso	569,393	45	250,436	10,416	2,067	19.8
Mato Grosso do Sul	442,099	40	198,142	14,060	2,060	14.7
Minas Gerais	3,066,879	227	821,631	11,275	1,048	9.3
Pará	832,355	56	79,584	2,125	488	23.0
Paraíba	430,538	40	99,891	5,207	716	13.8
Paraná	1,878,055	87	1,454,639	44,362	6,583	14.8
Pernambuco	977,721	82	289,921	9,155	1,109	12.1
Piauí	307,818	31	33,474	1,843	254	13.8
Rio de Janeiro	2,362,499	127	1,013,274	8,708	922	10.6
Rio Grande do Norte	379,473	38	36,130	2,307	195	8.5
Rio Grande do Sul	1,791,515	128	662,611	14,273	1,519	10.6
Rondônia	234,515	20	52050	6,221	921	14.8
Roraima	61,362	7	9,081	5,880	800	13.6
Santa Catarina	1,183,483	74	324,924	9,947	1,026	10.3
São Paulo	10,045,183	315	4,409,235	27,270	1,970	7.2
Sergipe	293,09	21	25,949	3,100	245	7.9
Tocantins	212,324	16	139,568	22,394	4,002	17.9
Total	31,424,197	1,930	12,295,072	14,098	1,697	12.0

Source: Authors' calculations, based on data from the Ministry of Labor.

almost 50 percent of the unemployed job seekers are high school graduates, fewer than 11 percent have any college education. Fifty-eight percent of the unemployed job seekers getting referrals are male, and 61 percent are considered nonwhite.

Brazil is well known for having wide regional cultural and economic variation, and this variation extends to the SINE system. Table 3 summarizes regional differences across Brazilian states when it comes to the provision of services in SINE offices. These heterogeneities suggest that differences across states should be considered in the process of estimating the impacts of SINE services.

Methodology

The purpose of this paper is to estimate the effects of SINE job interview referrals on labor market outcomes. That is, we analyze the effect of job interview referrals by SINE offices on the labor market outcomes of recipients relative to nonrecipients. However, simple differences of means between recipients and nonrecipients will not yield causal estimates of program effects because the characteristics of the two groups are likely to differ, owing to self-selection into SINE registration and services.

The evaluation problem is to compare workers who received a SINE job referral to their counterfactuals without a job referral. The challenge is to make sure the counterfactual is properly selected. SINE services match workers to vacancies based on a list of criteria, and this automated process with mediation by SINE staff might be more efficient than workers trying to find a job match by themselves.[17] However, we do not observe the outcome for service recipients had they *not* received the service—the ideal counterfactual. In this study, we use PSM to construct a counterfactual for the group getting referrals—the participant group—by selecting a group of registered workers who are not getting referrals but who have a similar pretreatment conditional probability of receiving a referral—the comparison group. We then estimate group mean effects, or the average treatment effect on the treated, as a difference in mean outcomes between these two groups. The individuals in the matched comparison group will be similar to the participants in terms of observed characteristics, except for the referral. The application of PSM requires satisfying the conditional independence and common support assumptions.[18]

17. The matching algorithm is based on occupation (up to seven occupations can be listed using the CBO, the Brazilian classification of professions), educational attainment, work, language skills, availability for traveling or staying away from home for long periods of time, and possession of a driver's license.

18. The assumption of conditional independence (selection on observables) requires that, conditional on a set of observed attributes, the distribution of the (counterfactual) nontreatment outcome in the treated group is the same as the (observed) distribution of the nontreatment outcome in the nontreated group. The common support assumption requires that all treated individuals have a counterpart in the nontreated population. This means that values of **X** in equation 1 are related to similar propensity scores in the treatment and control groups. For details, see Blundell and others (2004) and Heinrich, Maffioli, and Vázquez (2010).

The propensity scores used to balance characteristics between participant (referrals) and comparison (not referred) groups are estimated using the following probit model for each subgroup evaluated:

$$(1) \qquad P\left(D = 1 \middle| \mathbf{X}\right) = \phi\Big[\beta\mathbf{X} + \gamma\Big(Age + Job_{tenure} + \log Wage + Gender + Unemployment\_spell\Big)D_{region}\Big].$$

In this specification, we calculate the probability of being referred for a job interview, $P(D = 1|\mathbf{X})$, as a function of observable individual characteristics. Importantly, our data include successive monthly cohorts of participants and their counterfactuals between January 2012 and December 2016, and job interview referrals are measured on a year-month reference basis.[19] Using these monthly samples of participants and nonparticipants, we estimate sixty PSM models. That is, we estimate separate PSM models on each monthly data set of treated workers in our panel.[20] We follow the approach of Sianesi (2004), who estimates separate PSM models for each month in her panel data.[21] We use nearest-neighbor matching within the same state without replacement to create comparison groups.[22]

19. In other words, we count referrals and registrations in a given month only once. Workers who successfully get reemployed are removed from the sample.

20. For each subgroup analysis performed in the Results section, sixty PSM models were estimated, thus creating different common supports with a different number of observations.

21. Sianesi (2004) evaluates employment services in Sweden and develops this monthly subsample approach, because nearly every customer of the employment service gets at least one service at some point. Constructing monthly samples allows for program participants and nonparticipants in each month. Other job referrals in the same month or later months—or other services in later months—could be confounding factors in our evaluation design. Therefore, we assume that the distribution of those receiving subsequent employment and training services is balanced between referrals and comparison group members.

22. The use of the closest match minimizes the bias, as we guarantee the use of the most similar observation to construct the counterfactual (Heinrich, Maffioli, and Vázquez, 2010). In other words, the match uses the closest propensity score to match one worker in the treatment group to a worker in the comparison group. We used the nearest matching without replacement, meaning workers in the control group are used only once as a match. Matching without replacement performs well when many comparison units overlap with the treatment group (Dehejia and Wahba, 2002). There is a large availability of observations in the control group, and appendix B shows that treatment and control groups overlap. Thus matching without replacement is appropriate in our setting.

The term ϕ is the normal cumulative distribution function. The remaining observable individual characteristics in the vector **X** for the PSM are as follows: tenure of the last job before referral (in months), the logarithm of the average monthly wage on the last job, race (divided into five categories: indigenous, white, dark, yellow, and brown), age in the year of the matching, gender, educational attainment (divided into eleven categories), industrial sector (eighty-six CNAE categories at the two-digit level) and occupational group (forty-eight CBO categories at the two-digit-level) of the person's last job, and number of months unemployed.[23] In addition, as shown in equation 1, age, job tenure, wage, gender, and unemployment duration are interacted with region dummy variables.[24] Tenure in the last job before referral (months) and the logarithm of the average monthly wage at the last job were included in the PSM to reduce selection on unobservables, as these variables encompass information on unobservables (Heinrich, Maffioli, and Vázquez, 2010).

We construct control groups using the pool of workers who registered at a SINE office but were not referred for a job interview in a given month. This approach mitigates selection bias on unobservables, since workers who visit a SINE office might have self-selected and received a job interview referral because of unobservable characteristics, such as their level of self-motivation and general proactiveness.[25] Additionally, we require the common support condition to be met exactly.

After estimating propensity score models, the next step is to perform the matching and assess its quality. The literature suggests that observable

23. CNAE is the national classification of economic activities; CBO, the national classification of professions. Since the large number of observations allows, we also estimated an alternative PSM whereby individuals are matched with certainty on two characteristics: the number of months unemployed until matching and the workers' state of residence. Thus, each treated individual is matched with a nontreated individual from the same state—someone who also has the exact number of months unemployed until matching. These additional results are available on request. The strategy of matching on exact characteristics is used by Lechner (2002), who performs matching using propensity scores and matching exactly on sex, duration of unemployment, and native language.

24. Heinrich, Maffioli, and Vázquez (2010) suggest that in a scenario with a limited number of variables, to obtain a balance between treatment and control groups, interactions with an available variable can improve the matching. We interact the vector **X** with regions to achieve an improved matching model.

25. The information used in the PSM to construct control groups always comes from RAIS. While the main database used to compare the referred versus nonreferred individuals was the SINE, information from the RAIS was essential to calculate PSMs and measure the outcomes, since it allowed us to track the employment history of each job seeker.

TABLE 4. Selected Descriptive Statistics Pre- and Post-Matching

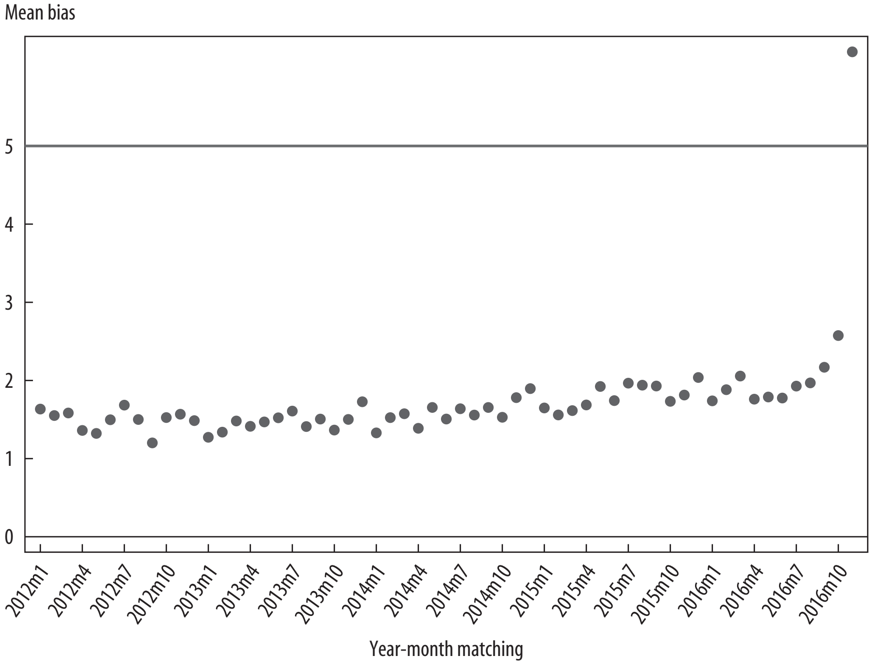
		Mean		Bias reduction			
Variable	*Sample*	*Treated*	*Control*	*Bias (%)*	*(%)*	*t test*	$P>\|t\|$
Male	Unmatched	0.549	0.583	7.050		20.064	0.00
	Matched	0.584	0.580	0.640	90.89	1.461	0.14
Age	Unmatched	31.474	32.831	12.580		36.922	0.00
	Matched	32.864	32.862	−0.270	97.78	−0.635	0.53
Tenure last job	Unmatched	24.073	15.594	−28.230		−94.025	0.00
	Matched	15.554	15.842	−1.126	96.00	−2.564	0.01
Mean wage last job (ln)	Unmatched	7.102	7.141	8.238		25.526	0.00
	Matched	7.143	7.144	−0.666	91.90	−1.517	0.13
White	Unmatched	0.445	0.460	2.914		8.263	0.00
	Matched	0.459	0.461	−0.151	94.81	−0.343	0.73
Elementary incomplete	Unmatched	0.029	0.031	1.518		4.260	0.00
	Matched	0.032	0.030	0.834	45.01	1.899	0.06
Elementary completed	Unmatched	0.031	0.030	−0.366		−1.042	0.30
	Matched	0.030	0.030	−0.347	−0.79	−0.790	0.43
Middle incomplete	Unmatched	0.081	0.085	1.550		4.371	0.00
	Matched	0.085	0.084	0.020	98.66	0.047	0.96
Middle completed	Unmatched	0.133	0.135	0.511		1.449	0.15
	Matched	0.135	0.151	−4.646	−808.64	−10.575	0.00
High school incomplete	Unmatched	0.165	0.126	−11.152		−32.558	0.00
	Matched	0.126	0.152	−7.481	32.68	−17.026	0.00
High school completed	Unmatched	0.478	0.542	12.467		35.405	0.00
	Matched	0.540	0.499	8.433	32.35	19.192	0.00
College incomplete	Unmatched	0.026	0.022	−2.659		−7.721	0.00
	Matched	0.022	0.017	3.591	−35.07	8.173	0.00
College completed	Unmatched	0.048	0.023	−13.518		−42.486	0.00
	Matched	0.023	0.027	−2.541	81.19	−5.784	0.00

Source: Authors' calculations, based on data from the Ministry of Labor.

Note: The treatment or participant group is made up of workers registered with SINE who received a referral for a job interview; the control or comparison group is made up of workers registered with SINE who did not receive a referral for a job interview in January 2016. The bias for a given variable is defined as the difference between the means of the treatment and control groups, scaled by the average variance.

characteristics should be balanced between the two groups after matching. As the matching is performed monthly, the balance in the means of basic obervable characteristics must be checked for each month. Table 4 shows the *t* tests for differences in means before and after the matching for certain characteristics in November 2016. The bias for a given variable is defined as the difference between the means of participant and comparison groups, scaled by the average variance. A bias reduction after matching is expected. The *t* tests indicate that before matching, the participant and comparison groups are sigificantly different on most observable characteristics, but after matching

FIGURE 1. Mean Standardized Bias between Participant and Comparison Groups Post-Matching



Source: Authors' calculations.
Notes: Each dot in the figure represents the mean standardized bias between control and treatment groups for each month of the data. The solid line marks the threshold of 5 percent final bias.

there are few significant differences. This suggests that the participant and nonparticipant matched samples are well balanced.

The matching does not necessarily need to yield complete balance on all exogenous variables to be satisfactory. We use the mean standardized bias to formally assess the quality of the PSM. If the matching process improves balance on observable characteristics between the participant and comparison groups, it is expected that the mean standardized bias between the two groups will be significantly reduced. According to empirical studies, a final bias below 5 percent after matching should be sufficient (Caliendo and Kopeinig, 2008). Figure 1 plots the value of the mean standardized bias calculated separately for each month. In this case, the bias maintains an average value of 1.7 after the matching, an indication of the good quality of the PSM.[26] An additional step to verify the matching quality is to examine the kernel density distribution graphs of the propensity score for the two groups before and after matching (see figures B1 and B2 in online appendix B).[27] These figures show that there is an overlap in the mean propensity scores and their distributions for the two groups after matching, suggesting that the PSM generates good matches.[28]

We use the participant and comparison groups constructed by PSM to measure impacts on the following labor market outcomes: probability of employment within three months, time from registration until employment, job tenure in the new job, and reemployment monthly earnings. As described earlier, to perform the matching, we restricted the database to workers who had lost their jobs prior to SINE job referral, which allowed us to calculate the pre- and post-matching variables. Details on the calculation of the resulting outcomes (pre- and post-treatment) are provided below.

Measuring SINE Impact on Labor Market Outcomes

Having used propensity score matching to construct counterfactual groups for workers who had a SINE job interview referral, which were validated by three tests, we use the participant and constructed comparison groups in the following difference-in-differences specification to estimate the impact of a job interview referral on labor market outcomes for worker i:

$$(2) \qquad Y_{it} = \phi + \alpha Treated_i + \gamma Post_{it} + \theta SINE_{it} + \beta \mathbf{X}_{it} + \mu_t + \varepsilon_{it},$$

where Y_{it} stands for one of the four outcome measures for individual i and time t. *Employment within three months of referral* establishes whether at the month of the matching the worker had gotten a job within three months of the referral. In the evaluation, this variable is always zero for the pre-matching

26. We also use the Rubin ratio test (see Rubin, 2001), and the results confirm the quality of the matching, as the ratio of variances of the propensity score and covariates from the treatment and comparison groups is close to 1.0, and it is between 0.5 and 2.0 for each of the sixty months (see figure B3 in the online appendix).

27. Supplementary material for this paper is available online at http://economia.lacea.org/contents.htm.

28. The PSM is conducted for each month of our panel, and the kernel densities present a similar pattern in every month. Monthly results are available on request.

period.[29] *Time until employment* is period of unemployment between jobs, calculated as the date of admission to the next job minus the date of separation from the previous job.[30] *Mean tenure* is the number of months in the reemployment job, and *reemployment wage* is the natural logarithm of the real wage on the reemployment job.[31]

The term ϕ captures all time-constant factors that affect the outcome. *Treated* is a dummy variable indicating whether the individual gets a SINE job referral or not, and *Post* takes the value of one after treatment. The variable *SINE* is the interaction between *Treated* and *Post*, whereas θ, the coefficient of interest, measures the difference in the outcome variable between the treated and control groups before and after receiving services from SINE. μ_t are the monthly dummy variables. The matrix **X** includes alternative education and sector variables for individual workers who are not included in the PSM.[32] We also include information on whether the worker is a beneficiary of unemployment insurance, dummy variables for the *n*th unemployment insurance payment, and the total number of referrals.[33] Standard errors for statistical inference are computed with clustering at the state level.[34]

29. To evaluate this outcome, we remove matches from September 2016 onward in order to leave only observations that are well defined (individuals who possess at least three months of information for this outcome).

30. Unemployment (nonformal employment) is calculated as the time between two jobs prior to the treatment. The calculation of the outcome time until employment requires information on two jobs prior to the job referral, generating a smaller number of observations for the regressions for this outcome. No further restrictions are imposed.

31. The data for mean tenure and reemployment wages require the observation of one job before and after matching to measure the outcomes; no further restrictions are imposed. In contrast to the method used to calculate the time until employment, the information on job tenure is observed in the record of employment prior to job search and does not need to be constructed from observing two jobs prior to the matching. Tenure of the reemployment job is computed as the difference between the job start and end dates.

32. Education is disaggregated into three categories: unskilled (from illiterate to completed primary school), semiskilled (partial or completed high school), and skilled (any tertiary education). The sector of the last job from the IBGE classification is aggregated in the following categories: agriculture, industry, services, trade, construction, and other.

33. These variables are included in the difference-in-differences estimations, as they were not available when the main bulk of the PSM was calculated. Alternative estimations including these variables in the PSM or difference-in-differences estimations, without the variables included in vector **X**, provide similar results.

34. We assume that the observations are independent across states as labor market institutions differ. For instance, even though the minimum wage is defined nationally, each state in Brazil can set its own minimum wage above the national minimum wage, which can influence labor market dynamics. Regarding the SINE service, even though SINE is coordinated at the central level, each state manages its own SINE network and might apply different resources and managerial procedures.

TABLE 5. Effect of SINE Job Interview Referrals on Labor Market Outcomes

Description	*Employment within three months*	*Time until employment (months)*	*Mean tenure (months)*	*Reemployment wage (log)*
Effect of SINE Standard error	0.200***	−0.452**	−3.533***	−0.058***
	(0.010)	(0.173)	(0.233)	(0.006)
No. observations	20,359,236	9,233,184	14,738,524	14,699,527

Source: Authors' calculations.
*$p < 0.10$; **$p < 0.05$; ***$p < 0.01$.
Note: The table shows the difference-in-differences estimates of the effect of SINE job referral on labor market outcomes. Standard errors clustered at the state level are in parentheses.

Results

The analysis seeks to measure the effect of job interview referrals on the probability of workers finding a job within three months of the referral, the time until employment, the mean tenure in the next job, and the reemployment wage. Impacts are computed by comparing outcomes of workers who received SINE job referrals to those of a matched comparison group of workers who were registered with SINE but did not get a job referral.[35]

The results show that the treatment increases the likelihood of finding a job within three months of the referral by 20.0 percentage points (see table 5). The probability of the control group participants finding a job within three months is 24 percent; thus a SINE interview referral nearly doubles their probability of finding a job within that time.[36] In addition, job seekers who are referred by SINE take less time (0.5 months less) to find a job than those who are not referred. This represents about a 6 percent reduction in the waiting time until they are able to secure a job, as in the control group the wait time is eight months, on average. However, SINE job referrals have a negative impact on the mean tenure in the next job found. On average, job tenure is reduced by 3.5 months, which equates to an 18 percent reduction in the average job tenure of 19.6 months found in the data.[37] Finally, being treated by SINE reduces wages by about 5.8 percent.

The result that a SINE job referral is associated with a wage reduction is consistent with Pignatti (2016) and Vera (2013) and may be due to stigmatization

35. Results using RAIS for control groups are very similar and are provided in online appendix C.

36. Online appendix C provides an indication of the size of the employment's system's impact on outcomes. For instance, 0.24 percent of workers in the control group obtained a job within three months after matching, and SINE increased this probability by 0.20 percentage points.

37. See footnote 12.

effects on SINE participants or a lack of capacity by the program to attract high-paying enterprises to the system.[38] Also, SINE job referrals promote faster reemployment, since SINE mainly lists low-wage jobs that have short tenure. This result is consistent with suggestions from job search and matching models that heterogeneous preferences for job amenities will be reflected in the distribution of reemployment wages and other attributes such as job durability (for example, McCall, 1994). The estimated effects are the average for the period of analysis, and because of the short job tenure and high worker turnover in the Brazilian labor market, the five-year time span is sufficient to provide results about how SINE affects labor market outcomes.[39] Subgroup analysis based on workers' characteristics is provided in the next section.

Demographic Subroup Analyses

Subgroup estimates reveal differences in the impacts of SINE services across groups of customers. These estimates help shape the strategy for providing services to workers with different characteristics. Our method for estimating subgroup impacts involves estimating a separate PSM for each subgroup category in each of the sixty months, using these to create matched-pair comparison groups for each subgroup category, and then estimating the effects of job referrals by difference-in-differences for each subgroup category.[40]

38. We used PSM to match firms that posted vacancies at SINE in 2015 and firms that did not. Matching variables were the proportion of males, proportion of white workers, average worker age, firm size, sector classification, and state of the firm. This exercise suggests that wages at a firm that posts vacancies at SINE are 140 Brazilian reais lower than wages at a similar firm that does not post vacancies at SINE. Other results indicating that SINE referrals decrease the time to reemployment but also reduce wages and time of employment need further investigation, as getting a job faster may be related to a worse quality of matching. Nevertheless, the overall data do not provide a clear correlation between time until employment and tenure/wage.

39. Table D1 in online appendix D provides separate estimates for the 2012 cohort as this cohort has a longer time span for the outcomes to materialize and thus mitigates for censored data, mainly for the time until employment and mean tenure outcomes. The results for the 2012 cohort are qualitatively similar and suggest a better performance for SINE referrals since the time until employment is further reduced and the negative impact on the average mean tenure is smaller.

40. The effects across groups are not directly compared with the overall effects as the difference-in-differences estimations and PSMs are conducted separately for each subgroup (for example, comparing women who get interview referrals to women who do not get interview referrals) to allow for the best matching and estimations against each control group. Alternative results for the full model, based on one general PSM, and estimations of subgroup effects in the same regression are available on request. Complete models are estimated for gender, education, age, race, and receipt of unemployment insurance. Estimating coefficients in the same regression allows for a better comparison across different groups and across different tests of the equality of coefficients; however, it provides poorer matching, as those treated in subgroups might be matched with a control who belongs to another subgroup.

Procedures for constructing samples to measure each of the four outcomes follow the same steps as listed in the Methodology section. Impact estimates for subgroups defined by characteristics of age, sex, race, and educational attainment are presented in table 6.

The general pattern of effect estimates on outcomes for each subgroup is similar to the full sample pattern of impact estimates presented in table 5: that is, they show a higher percentage of employment within three months of the job interview referral, fewer months until reemployment, fewer months of job tenure in the new job, and lower reemployment earnings. However, there are some significant differences in impact estimates between some subgroup categories.

By age group, the positive effects of SINE referrals on the time to find a job are smallest for the youngest workers (eighteen to twenty-four years of age). Indeed, the youngest group has a significantly smaller positive effect than all age groups.[41] The effect on shortening the time until reemployment is significantly greater for the oldest group (fifty-five to sixty-four years) and significantly smaller for the youngest group (eighteen to twenty-four); the estimated effects for the other age groups fall about in the middle of that range.[42] The effects on decreasing tenure in the new job grow steadily larger with age. These effects are significantly different among the five age groups, rising steadily from 2.096 fewer months in the youngest age group (eighteen to twenty-four) to 6.950 fewer months in the oldest age group (fifty-five to sixty-four years). Job referrals reduced reemployment wages the most for the younger prime-age workers (twenty-five to thirty-four), at a rate of 5.9 percent. This reduction is significantly larger than for the youngest workers (eighteen to twenty-four), who had a rate of 4.1 percent. Reemployment earnings reductions for the three older age groups declined with age, falling from 5.6 percent (thirty-five to forty-four) to 5.2 percent (forty-five to fifty-four), to 5.0 percent (fifty-five to sixty-four).

By gender, the impact of a SINE job interview referral had significantly better effects for men than for women on the probability of finding a job. For men, the increase in the probability of reemployment within three months is

41. The results for the fifty-five to sixty-four age group are influenced by retirement, as Brazil's average retirement age is fifty-six years for men and fifty-three years for women. Prior to 2019, a minimum number of years of contribution to the system provided eligibility for pensions, irrespective of age, because of legislation in place during the period analyzed in this paper (OECD, 2017).

42. Alternative results provided for the 2012 cohort suggest a clearer stronger effect on shortening the time until reemployment as age increases.

TABLE 6. Estimates of SINE Job Interview Referral Impacts, by Subgroup

Sample subgroup	*Employment within three months*	*Time until employment (months)*	*Mean tenure (months)*	*Reemployment wage (log)*
Age 18–24 years	0.226***	–2.330***	–2.096***	–0.041***
	(0.012)	(0.103)	(0.116)	(0.003)
N	3,928,116	1,761,790	2,657,300	2,649,949
Age 25–34 years	0.267***	–3.107***	–2.762***	–0.059***
	(0.008)	(0.108)	(0.240)	(0.006)
N	8,366,676	4,570,504	5,728,910	5,713,302
Age 35–44 years	0.265***	–3.185***	–3.398***	–0.056***
	(0.009)	(0.127)	(0.449)	(0.008)
N	4,808,100	2,431,800	3,041,026	3,032,629
Age 45–54 years	0.254***	–3.105***	–4.919***	–0.052***
	(0.009)	(0.152)	(0.584)	(0.009)
N	2,416,680	1,130,826	1,401,982	1,398,012
Age 55–64 years	0.242***	–3.884***	–6.950***	–0.050***
	(0.010)	(0.185)	(0.488)	(0.010)
N	779,760	337,192	391,184	390,046
Male	0.275***	–3.180***	–4.028***	–0.064***
	(0.009)	(0.094)	(0.365)	(0.009)
N	11,707,680	6,339,806	7,858,306	7,837,233
Female	0.238***	–3.836***	–4.213***	–0.065***
	(0.009)	(0.124)	(0.303)	(0.005)
N	8,678,488	3,684,396	5,363,858	5,348,523
White	0.260***	–3.750***	–4.503***	–0.078***
	(0.011)	(0.138)	(0.366)	(0.008)
N	9,585,256	4,642,246	6,250,658	6,232,846
Nonwhite	0.259***	–3.207***	–3.696***	–0.052***
	(0.007)	(0.099)	(0.287)	(0.006)
N	10,800,780	5,392,306	6,968,744	6,950,172
Unskilled	0.287***	–3.686***	–4.237***	–0.019**
	(0.010)	(0.184)	(0.485)	(0.008)
N	3,368,556	1,679,206	2,144,906	3,368,556
Semiskilled	0.254***	–3.400***	–3.952***	–0.061***
	(0.009)	(0.100)	(0.318)	(0.006)
N	16,202,160	7,965,430	10,577,488	10,549,066
Skilled	0.240***	–3.304***	–5.765***	–0.235***
	(0.011)	(0.162)	(0.399)	(0.014)
N	815,440	398,982	503,476	502,265

Source: Authors' calculations.

*$p < 0.10$; **$p < 0.05$; ***$p < 0.01$.

Note: The table shows the difference-in-differences estimates of the effect of SINE job referrals on labor market outcomes for demographic subgroups. Standard errors clustered at the state level are in parentheses.

larger, at 27 percentage points, versus 24 percentage points for women. On the other hand, a SINE referral reduces women's time until employment by 3.8 months, as opposed to 3.1 months for men. There were less appreciable differences between the genders in the reduction in reemployment job tenure or the reduction in reemployment earnings.

With respect to differences in impacts by race, SINE job referrals had generally better impacts for nonwhites than for whites. There was no difference by race in the impact on the probability of employment within three months and the time to reemployment was slightly more reduced for whites than for nonwhites. However, the reduction in new job tenure was bigger for whites, as was the reduction in reemployment wages. RAIS is an administrative database in which employers classify the race of employees according to subjective criteria. This can be particularly problematic in a country as diverse as Brazil. Paixão and others (2012) and Câmara (2015) present results showing discrepancies in data on race between the RAIS database, the IBGE National Household Sample Survey (PNAD), and the national census. The differences are significant, as RAIS presents a higher proportion of whites than PNAD and the census.[43] Using RAIS data, Cornwell, Rivera, and Schmutte (2017) show that when a worker changes jobs, the new employer might report a different race than the previous employer, and differences in race reporting are systematically associated with variation in wages. Thus our results by race must be interpreted with caution.

Only 10 percent of workers who seek SINE job search assistance have any tertiary education. While there is self-selection in the level of educational attainment, simple subgroup differences in impacts on employment outcomes by educational attainment help to inform decisions on program refinement. We grouped educational attainment into three categories: unskilled (from illiterate to completed primary school); semiskilled (partial or completed high school); and skilled (any tertiary education). Most job referrals (63 percent) went to semiskilled workers, while only 10 percent were in the skilled group. The magnitude of the effect of job referrals on the probability of finding a job within three months decreases signficantly as educational attainment increases. This means that in relative terms, SINE job referrals benefit less skilled job seekers the most. As for the other subgroup regressions, all education

43. Paixão and others (2012) show that in 2009, RAIS identified 61.2 percent of individuals as white, while PNAD identified 54.7 percent of workers as white. Câmara (2015) shows that in 2010, RAIS identified 60 percent of workers as white, whereas the 2010 census identified only 53 percent of workers as white. Race in the RAIS data is disaggregated into five categories (indigenous, white, dark, yellow, brown). For table 6, we divide the data into white and nonwhite.

categories see a big reduction in the time until reemployment as a result of a SINE job referral. The unskilled and semiskilled had the smallest reductions in reemployment job tenure, significantly smaller than for skilled job seekers. The impact on reemployment wages of a SINE job referral was significantly smaller for the unskilled (−1.9 percent) than for the semiskilled (−6.1 percent) and the skilled (−23.5 percent). The negative effect on the wages of the highly skilled might signal incapacity on the part of SINE to attract high-quality vacancies. As other researchers have found for other countries, our evidence suggests that SINE job referrals are particularly valuable for the unskilled, especially regarding the probability of finding a job and the reemployment wage.

Effects by Unemployment Insurance Recipiency and Unemployment Duration

The analysis based on unemployment insurance (UI) status is relevant because the effectiveness of the service for UI beneficiaries might be different, and there is evidence that access to UI affects incentives for formal employment. Tatsiramos (2014) points out that UI systems can increase reservation wage and lead to longer unemployment spells. However, UI benefits can provide the conditions for UI beneficiaries to increase the quality of the job found. Furthermore, Carvalho, Corbi, and Narita (2018), van Doornik, Schoenherr, and Skrastins (2018), and Cravo and others (2020) find that Brazil's formal sector workers who have access to UI have the ability and incentives to induce their own dismissal to some extent.

The long-term unemployed form an especially vulnerable group of applicants, defined as people who have been unemployed for more than twelve months. Results for this group go in the same direction as results for the full sample, but show differences in the magnitude of the effects (see table 7). The effect of SINE job referrals is stronger for this group in terms of the likelihood of finding a job within three months and the time it takes to get a job, which is 1.6 months shorter than for long-term unemployed who did not get a SINE job referral. Nevertheless, the negative impact on wages is more pronounced for long-term unemployment, as finding a job through a SINE job referral reduces wages by about 10 percent.

The results for the analysis based on unemployment status show heterogeneity in the impact of the labor intermediation process. In particular, unemployment insurance benefits may affect the results of the labor intermediation process, which has implications for unemployment spells and the quality of the job matching. While deeper investigation is warranted, SINE job referrals appear to be an effective means of reducing long-term unemployment.

TABLE 7. Effects of SINE Job Interview Referrals by UI Receipt and Unemployment Duration

Sample subgroup	*Employment within three months*	*Time until employment (months)*	*Mean tenure (months)*	*Reemployment wage (log)*
UI beneficiaries	0.207***	−2.533***	−2.795***	−0.029***
	(0.008)	(0.103)	(0.486)	(0.005)
N	2,157,364	1,123,086	1,666,510	1,663,046
Non-UI beneficiaries	0.227***	−3.131***	−2.754***	−0.055***
	(0.011)	(0.087)	(0.144)	(0.005)
N	11,483,120	5,808,344	7,532,858	7,510,053
Long-term unemployed	0.298***	−2.122***	−4.974***	−0.099***
	(0.009)	(0.094)	(0.503)	(0.011)
N	7,125,368	2,329,738	4,555,288	4,544,947

Source: Authors' calculations.

*$p < 0.10$; **$p < 0.05$; ***$p < 0.01$.

Note: The table shows the difference-in-differences estimates of the effect of SINE job referrals on labor market outcomes for subgroups of workers with different unemployment insurance (UI) status and experiencing long-term unemployment. Standard errors clustered at the state level are in parentheses.

Staff-Assisted versus Self-Service Job Referrals

Technology is changing the way in which public services are provided. Digital channels for labor intermediation have been adopted in many countries; these contribute to the effectiveness and efficiency of the public employment service. Nevertheless, little empirical evidence is available on how mobile technologies affect labor intermediation services and employment outcomes. Dammert, Galdo, and Galdo (2015) provide one exception, as they designed an experiment to assess the causal impacts of digital public labor market intermediation in Peru. The authors suggest that the use of digital technologies in the public labor intermediation system increases the probability of finding employment in the short term.

To contribute to knowledge on digital channels for labor intermediation, we investigate how online and face-to-face systems of service provision differ with respect to their effectiveness in placing job seekers in formal jobs and also with respect to the quality of the placements. This is an important aspect of intermediation services in many developed and developing economies, which have invested in developing online intermediation platforms as a means to increase coverage and reduce costs.

Table 8 shows the effect of SINE online referrals for one group versus the effect of using face-to-face referrals for a control group. The results show that the probability of getting a job within three months is not statistically different if the referral is online. However, the time until employment after the

TABLE 8. Effects of SINE Internet Referrals

Variable	*Employment within three months*	*Time until employment (months)*	*Mean tenure (months)*	*Reemployment wage (log)*
Effect from SINE (relative to control)	0.004	0.569***	0.540**	0.012**
	(0.010)	(0.135)	(0.248)	(0.005)
No. observations	283,872	185,924	198,560	198,079

Source: Authors' calculations.

*$p < 0.10$; **$p < 0.05$; ***$p < 0.01$.

Note: The table shows the difference-in-differences estimates of the effect of SINE job referral on labor market outcomes. The control group received face-to-face job interview referrals. Standard errors clustered at the state level are in parentheses. Results presented in this table should be interpreted with caution because of a shorter time span, as internet-based referrals only started in 2014.

referral is 0.6 months longer, suggesting that the face-to-face service is more effective. On the other hand, for those who obtain a job, the mean tenure is 0.5 months longer, and the reemployment wage is 1 percent higher. Thus our results suggest that face-to-face referrals are more effective than online service for obtaining employment faster, but job matching seems to be more efficient through online services as reemployment wages are higher and job tenure is longer.

Conclusion

This paper relies on the rich administrative records of SINE and RAIS to provide the first impact evaluation of SINE job interview referrals in Brazil on four labor market outcomes: the likelihood of reemployment, time to reemployment, job tenure in the new job, and the monthly reemployment wage rate. Using data from January 2012 to December 2016, we construct propensity scores matched pairs and compute difference-in-differences regressions to measure the impact of SINE on the four labor market outcomes. Overall, SINE job interview referrals increase the likelihood of reemployment in the first three months following referral and decrease the time to reemployment. Being referred by SINE has bigger effects for less skilled workers than it does for more highly skilled workers.

However, a job interview referral by SINE appears to reduce the job tenure in the new job and the monthly wage on that job. Stigmatization effects on program participants or the lack of capacity of the PES to attract high-quality job vacancy postings to the system might be contributing to these results.

The results of our study provide a clearer explanation of how SINE functions, and thus can contribute to the design of better labor market policy.

The heterogeneity of the system's impact on different subgroups suggests that providing specific support to each group of customers might improve the effectiveness of labor intermediation services. The use of technology for web-based job interview referrals contributes to the placement of workers, but face-to-face services have a greater impact on shortening the time until employment. Thus there appears to be room for technological improvement in the matching algorithm used for online services; such improvement could reduce the gap between face-to-face and remote services. A combination of services, provided at a SINE office as well as remotely, should be considered to increase the cost-effectiveness of the SINE network while maintaining its impact.

The heterogeneous effects of SINE on different groups of customers call for a more tailored approach to increase both the effectiveness and the efficiency of the intermediation services. Additional research is needed to understand the most cost-efficient combination of online and face-to-face services.

Note

We thank Eduardo Pontual Ribeiro, Daniel da Mata, Caio Piza, Rodrigo Quintana, Carlos Corseuil, Miguel Foguel, Paulo Jacinto, Aguinaldo Maciente, Hudson Torrent, Gustavo Alves Tillmann, Sinara Neves, Suely Barrozo Lopes, Ken Kline, Bassam Júnior, Mario Magalhães, Jociany Luz, Karla Carolina Calembo Marra, Wagner Rios, Mariana Almeida, José Ferreiro Espasandin, Diego Fernandes, and Viviane Cesario for their comments on earlier drafts.

References

Aedo, Cristian, and Ian Walker. 2012. *Skills for the 21st Century in Latin America and the Caribbean*. Washington, DC: World Bank.

Betcherman, Gordon, Karina Olivas, and Amit Dar. 2004. "Impacts of Active Labor Market Programs: New Evidence from Evaluations with Particular Attention to Developing and Transition Countries." Social Protection Discussion Paper 0402. Washington, DC: World Bank.

Blundell, Richard, Monica Costa Dias, Costas Meghir, and John Van Reenen. 2004. "Evaluating the Employment Impact of a Mandatory Job Search Program." *Journal of the European Economic Association* 2 (4): 569–606.

Borges, Márcio, Vinicius Lobo, and Miguel Foguel. 2017. "Análise da contribuição do programa de intermediação de mão de obra para os gastos com seguro-desemprego e para arrecadação previdenciária no período recente (2015–2016)." Technical Note 34. Brasília: Institute for Applied Economic Research (IPEA).

Caliendo, Marco, and Sabine Kopeinig. 2008. "Some Practical Guidance for the Implementation of Propensity Score Matching." *Journal of Economic Surveys* 22 (1): 31–72.

Câmara, Guilherme. 2015. "A qualidade das informações de cor ou raça na RAIS: Um estudo comparativo com o censo demográfico de 2010." Thesis, Federal University of Rio de Janeiro (UFRJ).

Carvalho, Cristiano C., Raphael Corbi, and Renata Narita. 2018. "Unintended Consequences of Unemployment Insurance: Evidence from Stricter Eligibility Criteria in Brazil." *Economics Letters* 162 (January): 157–61.

Cornwell, Christopher, Jason Rivera, and Ian M. Schmutte. 2017. "Wage Discrimination When Identity Is Subjective: Evidence from Changes in Employer-Reported Race." *Journal of Human Resources* 52 (3): 719–55.

Cravo, Túlio, Christopher O'Leary, Ana Cristina Sierra, and Leandro Veloso. 2020. "Heterogeneous Effects of Brazilian Unemployment Insurance Reforms on Layoffs." *Economics Letters* 197 (December): 1–4.

Crépon, Bruno, Esther Duflo, Marc Gurgand, and others. 2013. "Do Labor Market Policies Have Displacement Effects? Evidence from a Clustered Randomized Experiment." *Quarterly Journal of Economics* 128 (2): 531–80.

Dammert, Ana Cristina, José C. Galdo, and Virgilio Galdo. 2015. "Integrating Mobile Phone Technologies into Labor Market Intermediation: A Multi-Treatment Experimental Design." *IZA Journal of Labor and Development* 4 (1): 1–27.

Dehejia, Rajeev H., and Sadek Wahba. 2002. "Propensity-Score Matching Methods for Nonexperimental Causal Studies." *Review of Economic and Statistics* 84 (1): 151–61.

DIEESE (Departamento Intersindical de Estatística e Estudos Socioeconômicos). 2016. *Rotatividade no mercado de trabalho brasileiro*. São Paulo: DIEESE.

Heinrich, Carolyn, Alessandro Maffioli, and Gonzalo Vázquez. 2010. "A Primer for Applying Propensity Score Matching: Impact Evaluation Guidelines." Technical Note 161. Washington, DC: Inter-American Development Bank.

IPEA (Institute for Applied Economic Research). 2014. "Funcionamento e desenvolvimentos recentes da intermediação de mão de obra." *Boletim mercado de trabalho* 57. Brasília: IPEA.

———. 2020. "Políticas sociais: acompanhamento e análise." In *Trabalho e renda*, chap. 6. Brasília: IPEA.

Jacobson, Louis, and Ian Petta. 2000. "Measuring the Effect of Public Labor Exchange (PLX) Referrals and Placements in Washington and Oregon." OWS Occasional Paper 2000–06. Washington, DC: U.S. Department of Labor, Office of Workforce Security, Employment, and Training Administration.

Johnson, Terry R., Katherine P. Dickinson, and Richard W. West. 1985. "An Evaluation of the Impact of ES Referrals on Applicant Earnings." *Journal of Human Resources* 20 (1): 117–37.

Lechner, Michael. 2002. "Some Practical Issues in the Evaluation of Heterogeneous Labour Market Programmes by Matching Methods." *Journal of the Royal Statistical Society: Series A* 165 (1): 59–82.

Lobo, Vinicius, and Viviani Anze. 2016. "Duas diretrizes para a reestruturação do Sistema Nacional de Emprego." *Boletim mercado de trabalho* 61, October 2016. Brasília: IPEA.

McCall, Brian P. 1994. "The Effect of Job Heterogeneity on Reservation Wages." *International Economic Review* 35 (3): 773–91.

Michaelides, Marios, and Peter R. Mueser. 2018. "Are Reemployment Services Effective? Experimental Evidence from the Great Recession." *Journal of Policy Analysis and Management* 37 (3): 546–70.

OECD (Organization for Economic Cooperation and Development). 2017. "OECD Policy Memo: Pension Reform in Brazil." Paris: OECD General Secretariat, www.oecd.org/brazil/reforming-brazil-pension-system-april-2017-oecd-policy-memo.pdf.

Paixão, Marcelo, Irene Rossetto, Elisa Moncores, and Guilherme Câmara. 2012. "Investigação sobre qualidade da variável cor ou raça na RAIS através de um estudo comparativo com a PNAD do IBGE." Grupo de Trabalho (GT) 30. Paper presented at the 36th Encontro Anual da ANPOCS, Águas de Lindóia, Brazil, October 21–25.

Pignatti, Clemente. 2016. "Do Public Employment Services Improve Employment Outcomes? Evidence from Colombia." Research Department Working Paper 10. Geneva: International Labour Organization.

Ramos, Carlos A. 2002. "Las políticas del mercado de trabajo y su evaluación en Brasil." *Macroeconomía del desarrollo* 16. Santiago: United Nations Economic Commission for Latin America and the Caribbean.

Rubin, Donald B. 2001. "Using Propensity Scores to Help Design Observational Studies: Application to the Tobacco Litigation." *Health Services and Outcomes Research Methodology* 2 (3–4): 169–88.

Sianesi, Barbara. 2004. "An Evaluation of the Swedish System of Active Labor Market Programs in the 1990s." *Review of Economics and Statistics* 86 (1): 133–55.

Silva, Joana, Rita Almeida, and Victoria Strokova. 2015. *Sustaining Employment and Wage Gains in Brazil: A Skills and Jobs Agenda.* Washington, DC: World Bank.

Tatsiramos, Konstantinos. 2014. "Unemployment Benefits and Job Match Quality." *IZA World of Labor* 44 (July): 1–10.

Van Doornik, Bernardus, David Schoenherr, and Janis Skrastins. 2018. "Unemployment Insurance, Strategic Unemployment, and Firm-Worker Collusion." Working Paper 483. Brasília: Central Bank of Brazil.

Vera, Celia P. 2013. "A Quasi-Experimental Evaluation of the Public Employment Service in Peru." Zirve University.

Woltermann, Silke. 2002. "Job-Search Methods and Labor Market Transitions." University of Göttingen.

MAGDALENA BLANCO
Universidad de Montevideo, Uruguay

FELIPE CAROZZI
London School of Economics, UK

JOSÉ MARÍA CABRERA
Universidad de Montevideo, Uruguay

ALEJANDRO CID
Universidad de Montevideo, Uruguay

Mandatory Helmet Use and the Severity of Motorcycle Accidents: No Brainer?

ABSTRACT We study the impact of mandatory motorcycle helmet use laws on the severity and volume of road accidents in Uruguay by exploiting a change in the enforcement of the traffic law. Using a difference-in-differences design based on an unexpected change in policy, we report a sharp increase in helmet use and a five percentage point reduction in the incidence of serious or fatal motorcyclist accidents from a baseline of 11 percent. The benefits of helmet use are disproportionately borne by groups more likely to experience serious injuries, such as males or young drivers. We find no evidence of other responses in terms of either the volume or type of accident, suggesting that motorcyclists' behavior did not respond to differences in risk. We show that additional costs of enforcement for the relevant government agencies were negligible and estimate the health benefits of the policy.

JEL Codes: I12, I18, R41, H89

Keywords: Law enforcement, safety and accidents, helmet use

Road traffic accidents are the leading cause of death for children and young adults worldwide. According to the World Health Organization, 1.35 million people die yearly in road accidents. The associated costs are estimated to account for roughly 3 percent of gross domestic product (GDP) in most economies. These costs are particularly high in the case of low- and middle-income countries, which register 93 percent of deaths (WHO, 2018). In an effort to curb the substantial human and material costs imposed by road traffic accidents, countries have implemented a panoply of different regulations, from mandatory seat belt and helmet use laws to vehicle speed limits. For several decades, economists have studied the effectiveness of

ACKNOWLEDGMENTS We thank members of the board of UNASEV (Unidad Nacional de Seguridad Vial), Sergio Urzúa, and Karen Macours for useful comments and suggestions. We are very grateful to Mariana Leguisamo for outstanding research assistance. A previous draft of this paper was circulated under the title "Effects of Motorcycle Helmet Laws on Fatalities' Prevention: An Impact Evaluation."

seat belt use laws, in particular, because they can in theory modify the actual and perceived risks of driver behavior. In turn, this could hypothetically induce unexpected changes in accidents that may render regulation ineffective or counterproductive. This is known as the Peltzman hypothesis, after a study showing evidence of increases in pedestrian accidents as a result of seat belt regulation in the United States (Peltzman, 1975). While evidence in support of the Peltzman hypothesis has been elusive in recent studies of the consequences of seat belt use, examples of inadvertent consequences of protection gear have been documented in other activities.[1]

In this paper, we study the impact of a change in the enforcement of mandatory helmet use regulation in Uruguay on the severity and volume of road accidents involving motorcyclists and other road users. Mandatory helmet use laws for motorcyclists are common but not universal, and enforcement varies substantially between nations, with widespread enforcement issues in middle- and low-income countries. The potential effects of helmet use on the perceived consequences of speedy driving and other forms of risk taking are similar to those hypothesized in the case of seat belts. Yet there is limited evidence in the economics literature on the direct and indirect impact of helmet use enforcement on injury rates for motorcyclists. By using detailed administrative data on all reported road accidents in Uruguay, we can estimate these effects and study the impact of mandatory helmet use on the volume and severity of accidents taking place, both for motorcycles and for other vehicles.

Our empirical strategy is based on quasi-experimental variation in enforcement induced by changes in national laws in Uruguay. Mandatory helmet use was introduced in 2007 as part of the National Traffic Law, yet two departments—Uruguay is divided into nineteen territorial jurisdictions called departments—refused to enforce this regulation. This situation changed when the Misdemeanors Act was passed by parliament in 2013. As a consequence of this act, the department of Soriano started to enforce helmet use for motorcycle drivers and passengers. This induced an arguably exogenous change in enforcement that can be exploited for the purpose of our analysis.

We document the effect of this change in enforcement on the volume, type, and severity of road accidents in the two years after 2013. Our findings indicate a substantial reduction in the severity of motorcycle accidents, with estimates suggesting that helmet use leads to a five percentage point reduction in serious

1. For example, Chong and Restrepo (2017) study the effect of protective gear in ice hockey on player behavior. Pope and Tollison (2010) find increased on-track accidents in NASCAR as a result of the introduction of new safety regulations.

and fatal accidents (from a baseline probability of 11 percent) and a similar increase in the fraction of accidents resulting in minor injuries. This effect is of similar magnitude to that observed for the reduction in serious accidents induced by seat belt use reported elsewhere.[2] Despite this large magnitude, and in contrast to the implications of the Peltzman hypothesis, we find no evidence of risk compensation by drivers. Neither accident volumes nor the types of accident taking place change as a result of the increase in helmet use.

Using our coefficients in combination with estimates of hospitalization costs in the country and the value of statistical life, we can obtain a rough estimate of the health benefits resulting from enforcement of the helmet use law. By comparing these with motorcycle registration numbers, we also compute the nuisance cost of helmet use that would be required to offset the health benefits of this policy. Finally, we document differences in the effectiveness of helmet use on accident severity for different subpopulations and report that helmets appear to be more effective at reducing accident severity for the subpopulations more at risk of injury, such as males, young drivers, or victims of accidents taking place at night.

A small set of studies in economics have looked specifically at the effects of helmet use in traffic accidents.[3] Perhaps the closest to our work is Dee (2009), which provides estimates of the effect of the introduction/removal of helmet use laws in U.S. states on fatalities, using a panel specification.[4] Total fatality effects are meant to incorporate the direct effect of helmet use plus potential compensating behavioral adjustments by drivers. Dickert-Conlin, Elder, and Moore (2011) find evidence of increased availability of organ donations by deceased motorcyclists in U.S. states that repeal mandatory use laws. Carpenter and Stehr (2011) find that the introduction of mandatory bicycle helmet use laws for the young results in reduced fatalities. They also report a substantial reduction in cycling.

Our paper contributes to this literature by testing for the effect of helmet use on accidents and injuries in a context in which the change in enforcement is induced by a national reform and arguably affects helmet use only. Perhaps more important, we provide the first causal estimates of the effect

2. Centers for Disease Control and Prevention (CDC), "Policy Impacts: Seat Belts," December 2010 (www.cdc.gov/transportationsafety/seatbeltbrief/index.html).

3. Studies in the fields of accident prevention and medicine also look at this question using a variety of empirical methods. Some recent examples include Houston and Richardson (2008), Peng and others (2017), Olsen and others (2016), and Lee (2018).

4. Dee (2009) also provides complementary results using a within-vehicle specification similar in spirit to the analysis in Evans (1986).

of helmet use on injury severity outside the United States. This is particularly important insofar as enforcement issues are especially acute in low- and middle-income countries.

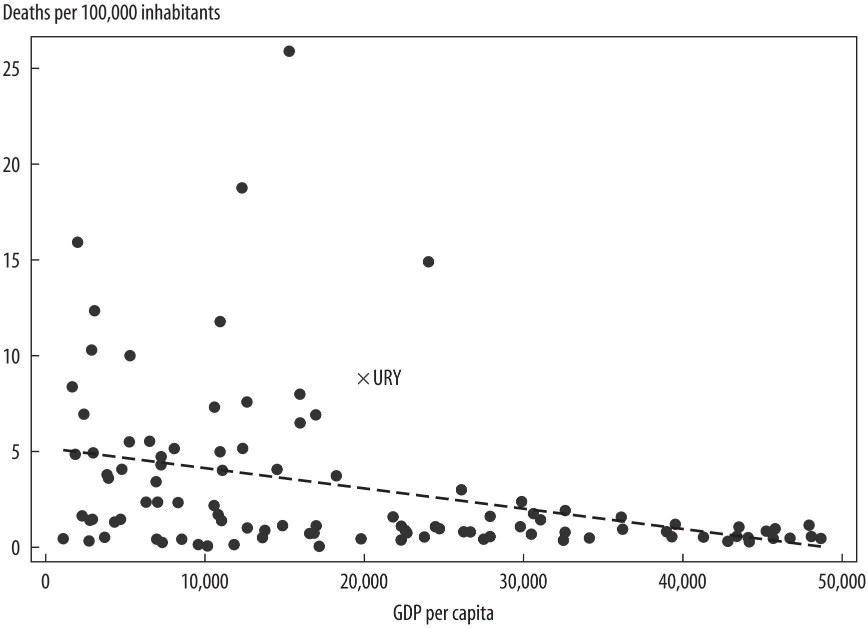
Our paper also relates to previous studies in economics estimating the impacts of seat belt use on health outcomes for drivers or nondrivers. Motivated by the work of Peltzman (1975), Loeb (1995) uses time-series data for Texas to study the effect of seat belt use laws on the fraction of accidents resulting in serious injuries. Cohen and Einav (2003) and Carpenter and Stehr (2008) improve the empirical strategy by exploiting a U.S. state panel. They respectively study the impact of seat belt laws on fatalities and injuries for vehicle occupants and nonoccupants. While we also exploit longitudinal variation by jurisdiction to estimate our effects of interest, there are important differences relative to these studies. In particular, we look at mandatory helmet use instead of seat belt use, and we use administrative data on individual accidents to investigate effects on the types of accident taking place. More important, we have information on injury type, which allows us to document impacts on serious and minor injuries and changes in composition between them.

Finally, our paper relates more broadly to the literature on policy solutions to the problem of road traffic accidents. Van Benthem (2015) uses historical changes in speed limits in the United States to obtain optimal limits, incorporating the impact of accidents and other factors (for example, air pollution). Hansen (2015) uses regression-discontinuity methods to study the impact of punishment for driving under the influence on recidivism. In an exception to the largely U.S.-centered literature, Aney and Ho (2019) study the impact of the Chinese Road Traffic Safety Law on the volume and severity of accidents and on fatalities. Our paper adds credible estimates of the effect of our policy of interest to the economics literature on policy solutions to traffic problems in the developing world.

Background and Data

Road traffic accidents are the leading cause of death for children and young adults aged five to twenty-nine years worldwide. The burden of road traffic injuries and deaths is disproportionately borne by vulnerable road users and those living in low- and middle-income countries, where the growing number of deaths is fueled by increases in transport motorization. Between 2013 and 2016, all low-income countries experienced an increase in the number of road traffic deaths (WHO, 2018). Despite the heavy costs imposed by road accidents,

FIGURE 1. Motorcycle Fatalities and Economic Development



Sources: World Health Organization and World Bank.
Notes: Sample corresponds to countries with GDP per capita of less than USD 50,000 (measured in 2013 dollars, at purchasing power parity). Country marked with a cross corresponds to Uruguay. Dashed line corresponds to a linear regression estimated over the scatter plot.

many countries still lack funded strategies, lead agencies, and adequate enforcement of existing traffic regulation.

Globally, those using motorized two- and three-wheelers—mainly motorcycle riders—represent 28 percent of all traffic-related deaths. The heavy burden of deaths borne by these road users is, at least in part, a result of them being less physically protected than car occupants. This additional risk for motorcycle users also affects the distribution of traffic-related deaths worldwide, as motorcycle use is generally more prevalent in developing countries.[5] Figure 1 shows a negative relationship between fatalities in motorcycle accidents

5. According to the 2014 Spring Pew Global Attitudes Survey, motorcycle ownership rates are regularly above 50 percent in developing East Asian economies, but less than 30 percent in developed countries.

and GDP per capita.[6] Our empirical analysis below focuses on Uruguay, which shows one of the worst rates in motorcycle accidents relative to its income level.

Tackling road safety problems in a context of increasing motorization is an important challenge for many developing economies. Even if adequate regulations are in place, these may be ineffective without the resources to ensure they can be successfully enforced. For example, in most countries helmet use is formally mandatory for motorbike drivers and passengers. Yet these regulations often coexist with low use rates: Argentina, Bolivia, Iran, Peru, and Uganda all have mandatory helmet use laws, but in these countries over 30 percent of drivers, and roughly 60 percent of passengers, do not wear helmets (WHO, 2018). The situation is often worse: in India and China, helmets are used by only 30 percent and 20 percent of drivers, respectively. Both countries have had mandatory helmet laws for over a decade.

That is not to say that mandatory helmet laws are universal. In the United States, many states require helmet use only for young riders (for example, under the age of twenty). The states of Illinois, Iowa, and New Hampshire do not require helmet use at all. In many of the countries that do have mandatory helmet laws, these laws do not specify standards for those helmets.

Helmet Use and Motorbike Accidents

When a motorcycle is involved in a collision, the rider is often thrown from the vehicle. In this event, a motorcyclist who is wearing a helmet has a lower risk of suffering traumatic brain injuries. There are typically three reasons for this. First, the helmet cushions the impact and therefore reduces the deceleration of the skull, which limits the speed of the impact between the brain and the skull. Second, a helmet spreads the force of the impact over a greater surface area so that it is not concentrated on a small area of the skull. Finally, helmets act as a mechanical barrier between the head and the object.

These three functions are met by combining the properties of three basic components of the helmet. The shell is the strong outer surface that distributes the impact over a large surface area. The impact-absorbing liner is the soft foam-and-cloth layer that sits next to the head. It helps keep the head comfortable and the helmet fitting snugly. Finally, the retention system or chin strap is the mechanism that keeps the helmet on the head in a crash.

6. Detailed information on per capita GDP and motorbike fatality rates by country can be found in table A1 in the online appendix (available at http://economia.lacea.org/contents.htm).

In the event of an accident, bikers who do not wear helmets generate additional hospitalization costs by needing a greater number of medical and surgical interventions and longer recovery times. The disability that often results from these head injuries leads to additional individual and social costs (WHO, 2006).

Natural Experiment

In November 2007, the Uruguayan parliament approved a new National Traffic Law (Law 18,191) that required mandatory helmet use for motorcyclists in all nineteen departments of the country.[7] However, the departments of Soriano and Cerro Largo decided not to monitor the use of helmets—effectively ignoring this aspect of the law. The local governments of both departments were able to sustain differential enforcement because the Uruguayan constitution devolves transit control to the departmental jurisdiction.[8] The refusal to enforce mandatory helmet use was partly based on electoral considerations, featuring prominently among the electoral promises in both departments. Both governors (*intendentes*) continued to promote the enforcement of speed limits and other elements of the national traffic laws

Perhaps as a result of the lack of enforcement, in early 2013—when our sample period starts—both departments had substantially lower reported rates of helmet use than other parts of the country. The percentage of motorcycle accidents in which the biker was wearing a helmet was 7.9 percent and 21.2 percent for Soriano and Cerro Largo, respectively. The average for other departments stood at roughly 75 percent. Moreover, helmet use was particularly low in Mercedes (the capital city of Soriano) and Melo (the capital city of Cerro Largo)—respectively 3.1 percent and 5.7 percent.[9]

In August 2013, parliament approved the Misdemeanors Act (Law 19,120), which includes an article establishing a specific punishment for motorcyclists not using a helmet, consisting of community work. In the months after the Misdemeanors Act was approved, the governor of Soriano informed citizens that the department would start enforcing mandatory helmet use.

7. A map of Uruguayan departments that includes the percentage of helmet use can be found in figure A1 in the online appendix.

8. Traffic inspectors are under the authority of local departmental governments and control traffic in urbanized areas. The national traffic police (*policía nacional de tránsito*) operates under the authority of the national government and focuses its attention on controlling traffic along national roads.

9. These cities have comparable numbers of registered motorcycles and automobiles per capita and similar helmet usage figures before 2013 (see table A2 in the online appendix).

"The Misdemeanors Act forced my hand," he stated in a press interview. "The local police chief asked me what to do, because if they saw someone not wearing a helmet they would have to proceed."[10] On November 1, 2013, the municipality of Soriano started monitoring motorcyclists. The department of Cerro Largo remained steadfast in its position, with the local government insisting on its jurisdictional priority. Cerro Largo does not, to this day, require helmet use for motorcyclists.

Two key assumptions are required to interpret the change in enforcement of the helmet use laws in Soriano as a natural experiment. The first assumption is that this change in policy is not correlated with previous or expected changes in helmet use or the volume and type of accidents in Soriano itself. We think this is a reasonable assumption in our context. The change in policy largely coincided with the approval of the Misdemeanors Act by the national parliament, and the governor specifically cited this approval as motivating the decision. The Misdemeanors Act was a substantial change to *national* legislation and was not itself a response to the traffic policy decisions of Soriano or Cerro Largo. Importantly, changes in the existing or expected severity of accidents are not mentioned as prompting the shift in policy.

The second assumption is that the change in mandatory helmet use did not come with other differential changes in local traffic policy. During this period, other traffic regulations in Soriano—on speed limits or drunk driving—were enforced regularly, which was often explicitly mentioned by the governor of Soriano before 2013 when defending his decision not to enforce the helmet laws.[11] According to administrative data on fines, the average number of fines issued by the Soriano traffic department before and after the policy change was stable.[12] This helps us to interpret systematic variation in the volume and

10. Emiliano Zecca, "Besozzi: No exigir el casco 'pudo ser un error,'" *Portal 180*, November 7, 2014 (www.180.com.uy/articulo/51928_Besozzi-No-exigir-el-casco-pudo-ser-un-error).

11. In statements to an Uruguayan news website, the Soriano governor declared, "We were betting on controlling drunk driving and speeding. We were strong with those [regulations] because 85 percent of accidents happened under the effect of alcohol or drugs or while speeding" (Emiliano Zecca, "Besozzi: No exigir el casco 'pudo ser un error,'" *Portal 180*, November 7, 2014 [www.180.com.uy/articulo/51928_Besozzi-No-exigir-el-casco-pudo-ser-un-error]).

12. This result is based on data from SUCIVE (*Sistema Único de Cobro de Ingresos Vehiculares*). The database has every fine for traffic offenses imposed in Soriano from January 2013 to December 2015. This encompasses 36,686 fines for motorcycles and 9,315 fines for cars. Figure A2 in the online appendix shows that the activity of traffic inspectors (reflected in the number of fines imposed on drivers) is not systematically different in the years before and after treatment. The difference in the average monthly number of fines to motorcyclists between periods is not statistically different from zero ($p = 0.66$).

type of accidents in Soriano relative to other departments as a plausible outcome of helmet use policy alone.

Data

We employ data drawn mainly from the UNASEV database.[13] This includes detailed information about the universe of accidents recorded by the police authorities, including the date, time, and location of each accident. The database includes information about the people involved in the accident, such as age, gender, role (whether the person was a passenger or a driver), consequence of the accident (death, serious injury, minor injury, or unharmed), and helmet or seat belt use as applicable. Locations in the original data set are reported with the latitude and longitude of each accident. We use location information to obtain the locality or town of each accident.

While the police report is filed by the officers who intervene in the accident, the health consequences of the accident are recorded by medical service personnel. They are responsible for identifying whether the person is slightly or seriously injured, with the difference depending on whether one or more vital organs are compromised. Death is registered to have happened as a consequence of an accident if it occurred either at the time of the accident or at the medical center within thirty days of the accident. During the period under consideration—from 2013 to 2015—203,725 people were involved in traffic accidents in Uruguay. Excluding pedestrians and accidents with missing location information, we have 149,873 observations in our database. Roughly 40 percent of those observations involved motorbikes. Twelve out of 100 people suffering motorbike accidents were seriously injured or killed, more than double the rate observed for other vehicles.[14] In the capitals of Soriano and Cerro Largo—Mercedes and Melo—3,378 persons suffered motorbike accidents in this period.

Table 1 shows the descriptive statistics for all reported motorbike accidents between 2013 and 2015, splitting the sample by helmet usage. Wearing a helmet is associated with a significantly lower probability of being seriously injured in motorcycle accidents, with riders wearing a helmet facing a 3.8 percentage point lower probability of being seriously injured or killed. This figure does not account for the potential endogeneity of helmet use. Motorcyclists make

13. National Road Safety Agency (Unidad Nacional de Seguridad Vial).
14. See table A3 in the online appendix.

TABLE 1. Descriptive Statistics: Motorbike Accidents by Helmet Use, 2013–15

	No helmet			*Helmet*			
Variable	*Mean*	*Std. dev.*	*N*	*Mean*	*Std. dev.*	*N*	*Mean differences*
Serious injury or death	0.14	(0.34)	14,283	0.10	(0.30)	46,406	0.037***
Slight injury	0.60	(0.49)	14,283	0.69	(0.46)	46,406	−0.095***
Unharmed	0.27	(0.44)	14,283	0.21	(0.41)	46,406	0.058***
Male	0.75	(0.44)	14,225	0.68	(0.47)	46,301	0.067***
Age	26.87	(13.53)	12,520	31.77	(14.04)	44,790	−4.900***
At night	0.32	(0.47)	14,283	0.27	(0.44)	46,406	0.052***

Source: Data from *Unidad Nacional de Seguridad Vial* (UNASEV), Uruguay.
***$p < 0.01$.

several decisions when riding their motorcycle: how fast to go, whether to respect traffic signs, whether to drive under the effects of alcohol or drugs, and whether to wear a helmet. Thus helmet usage is an endogenous choice variable. Riders who decide to use a helmet self-select themselves into this group, so there can be observable and unobservable factors that confound the use of a helmet and the severity of an accident. For example, table 1 shows that unhelmeted riders are disproportionately young, male, and riding at night. In the next sections of the paper, we estimate the *causal* effect of using a helmet on the probability of serious injuries and fatalities.

In our difference-in-differences analysis below, we compare the evolution of accident volumes, helmet use, and health outcomes in Soriano with the rest of the country. Table 2 presents a series of pretreatment descriptive variables for four relevant groups. Soriano has a population of roughly 83,000 inhabitants. According to the 2011 census, there were 140 cars and 302 motorcycles per 1,000 inhabitants in this department. The population and the number of cars per capita are lower in this department than in the average department in Uruguay. Conversely, Soriano has a large number of motorcycles per capita relative to the rest of the country. A substantial fraction of these differences can be attributed to Uruguay's capital, Montevideo. In our robustness checks, we verify that the main results of this paper are robust to excluding Montevideo from the comparison group in our difference-in-differences sample. With regard to accidents and motorcycle accidents in particular, both accident severity and total accidents per capita are fairly similar between Soriano and other departments.

Our accident data include only *reported* accidents. We expect the coverage of our data to be reasonably comprehensive, particularly for accidents in which

TABLE 2. Descriptive Variables by Department

Variable	*Soriano*	*Other departments*	*Other depts. (excl. Montevideo)*	*Cerro Largo*
Census 2011				
Mean population	82,594	177,960	110,855	84,698
Cars per 1,000	140	137	140	124
Motorbikes per 1,000	302	148	217	258
Accidents, UNASEV 2013				
All accidents				
Total per 1,000	9.3	16.3	15.0	15.8
Serious injury (%)	5.7	4.8	5.0	3.7
Minor injury (%)	36.3	34.6	35.8	37.3
Motorbike accidents				
Total per 1,000	5.6	4.8	5.5	6.0
Serious injury (%)	9.0	9.9	9.1	6.9
Minor injury (%)	56.6	69.0	68.5	70.4
Helmet (%)	6.5	78.8	75.5	20.1

the participants were injured or the vehicles were damaged.[15] That said, some accidents are surely missing from the UNASEV source. Therefore, we work with a selected sample of the total population of drivers; we cannot observe or document the helmet use of riders who were not involved in accidents or accidents that were not reported. This has two significant implications for our empirical analysis. The first is that cross-sectional differences in accident volumes and in the type of accident taking place may induce some degree of endogenous selection. Here is where the change in policy allows us to devise an empirical strategy that avoids this issue. The second implication relates to the interpretation of our findings. Our estimates of the impact of helmet use on the probability of having a serious accident are made relative to the population of bikers involved in a reported accident. We believe this is the population of interest from a policy perspective, particularly because we do not find an impact of helmet use on the volume of accidents. Nevertheless, the resulting estimates would be slightly lower in absolute terms if taken over the (unobservable) population of all accidents.

15. According to Law 18,191 (*Ley Nacional sobre Tránsito y Seguridad Vial*), all accidents resulting in personal or material damages must be reported to the relevant authorities. Third-party insurance is mandatory in Uruguay, and the associated payment can take place only if the accident was reported.

Empirical Analysis

Our empirical analysis has three main goals. The first is to evaluate the consequences of the change in enforcement of the mandatory helmet law, identifying effects on helmet use and the severity of road accidents. The second objective is to estimate the effect of helmet use itself on accident severity, using the policy change as a source of exogenous variation. Finally, we want to document any other noticeable changes in driving behavior resulting from the change in policy. We tackle these objectives by exploiting the abrupt change in enforcement of helmet use in the department of Soriano in November 2013. We do so in the context of a difference-in-differences framework where the evolution of accident volumes, helmet use, and accident severity in this department is compared with that of other locations in the country. The resulting difference-in-differences coefficients can be interpreted as an average treatment effect of the change in policy under the typical parallel trends assumption.

In addition to studying the impact of the change in enforcement on helmet use and accidents, we use our data to explore the heterogeneous impact of helmet use on different types of driver.

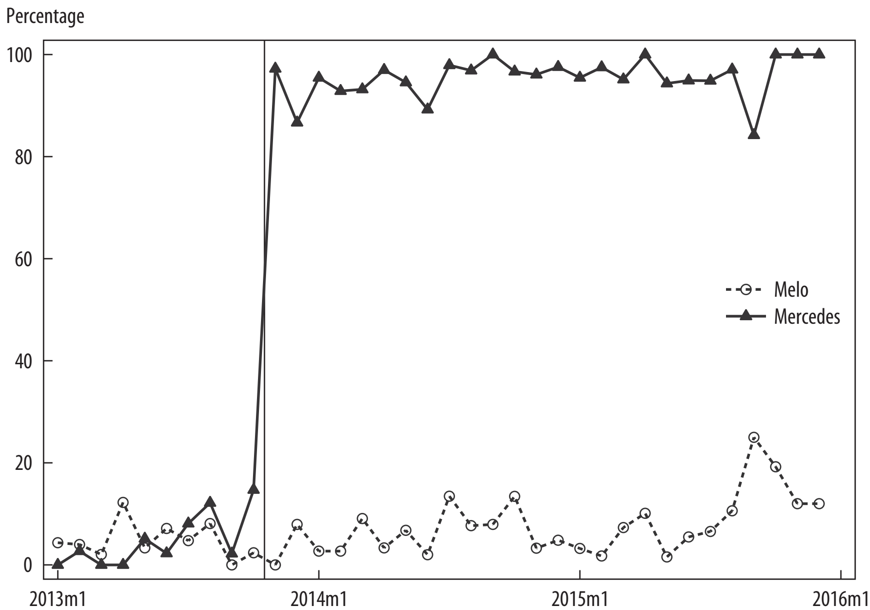
Illustration: Mercedes and Melo

In early 2013, the cities of Mercedes and Melo were the only department capitals in the country where municipal traffic inspectors did not enforce the helmet use law. As discussed above, Mercedes started enforcing that law in November 2013. To provide an initial illustration of the effects of the policy change, we report two event-study graphs comparing helmet use and the severity of motorbike accidents for both cities in figures 2 and 3.

Figure 2 plots the evolution of the percentage of people involved in a motorcycle accident who were reportedly wearing a helmet, for both cities. We use this variable as a proxy for helmet use. The initial levels of helmet use are remarkably low in both locations, oscillating under 10 percent. In November 2013, the rate of helmet use jumps to almost 100 percent in Mercedes, while the figures for Melo remain very low. This difference is sustained throughout the next two years and indicates that the change in enforcement prompted a persistent increase in helmet use in the city of Mercedes.

The evolution of the fraction of motorcyclists involved in accidents that experience serious or fatal injuries for both cities is reported in figure 3. We report three-month moving averages to smooth out some of the short-run fluctuations, but avoid smoothing between periods around November 2013.

FIGURE 2. Helmet Use in Mercedes and Melo

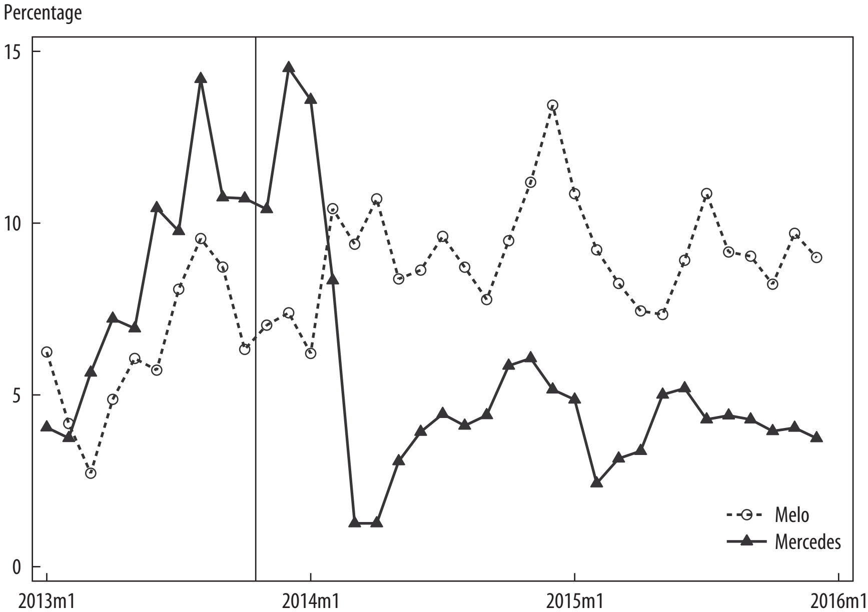


Source: Authors' calculations, based on data from UNASEV.
Note: Helmet usage is measured as the percentage of all motorbike accidents in which the driver was wearing a helmet. Vertical line corresponds to November 2013.

Before the change in enforcement, the fraction of serious accidents for both cities evolve in parallel with an upward trend, with the level being consistently higher in Mercedes. In the months before November 2013, the fraction of motorbike accidents resulting in serious injury in this city oscillated around 10 percent. Five months after the policy was introduced, serious injuries occurred in only 2 percent of motorbike accidents. Between late 2014 and 2015, the figure would recover to a level of around 4 percent. In this period, the rate of serious injury in Melo was twice as high as the rate in Mercedes. The fact that this divergence broadly coincides with the change in policy indicates that the increase in enforcement resulted in reduced injuries for bikers.

In the figure, the decline in serious accidents in Mercedes does not occur immediately after the change in enforcement, but rather takes about five months to materialize. In the first three months after the introduction, there is an apparent increase in the ratio of serious injuries. Given the changes reported in figure 2, we know this transition is not induced by a slow and progressive

FIGURE 3. Serious Injuries and Fatalities in Mercedes and Melo



Source: Authors' calculations, based on data from UNASEV.
Notes: Serious injury or fatality is defined as a percentage of the number of motorcycle accidents in each city. Vertical line corresponds to November 2013. The series represent three-month moving averages, where averages are taken without crossing the vertical line.

change in helmet use. A closer look at the raw data reveals that this was largely motivated by an abnormally high rate in January 2014, which resulted from a relatively low number of motorcycle accidents (thirty) coupled with a relatively large number of serious injuries (eight). Given the low numbers involved in that month, we do not interpret this spike as being an outcome of the policy.

Difference-in-Differences Strategy

To estimate the size of the effects of the change in enforcement of the helmet laws in Soriano, we use data for the universe of motorcycle accidents in all the country's localities in a difference-in-differences specification. We can thus incorporate data from all the towns and villages affected by the policy in the treatment group, while the comparison group is composed of all other towns in the country. The objective of the exercise is to obtain an average treatment effect that can be used to evaluate the benefits associated with the policy, as well as to identify potential unintended consequences.

Before estimating the effect of the policy, we use a locality-month panel to estimate whether the parallel trend assumption is reasonable in this context and whether the volume and type of accidents were affected by the policy. The first exercise is necessary to give causal interpretation to the difference-in-differences estimates below; the second, to narrow down the potential mechanisms relating helmet use to the change in accidents.

Our data set on road accidents starts in January 2013, so we have ten months to test for differences in pre-trends between the treatment towns in Soriano and comparison towns throughout the country. With these ten months of data, we use our town-month panel and estimate the following specification:

$$Y_{jt} = \alpha_j + \delta_t + \eta Post_t T_j + \varepsilon_{jt}, \tag{1}$$

where Y_{jt} represents the outcome variable in town j and month t, α_j represents a town fixed effect, and δ_t is a set of month-year dummy variables. The coefficient η multiplies an interaction of a treatment dummy variable T_j, which takes a value of one for the localities of Soriano, and $Post_t$, a dummy variable taking a value of one between June and October 2013.[16] We cluster standard errors at the locality (town) level and consider alternative methods for inference below.

A value of η statistically different from zero indicates that there were differences in pre-trends of the dependent variable between treatment and comparison groups before the policy change. We consider a set of different outcomes to detect trends both in our main variables of interest (namely, helmet use and accident severity) and in other correlates (such as the driver's age and gender and when and where the accident took place). Results are reported in table 3. We find no evidence of statistically significant differential pre-trends in any of the main variables of interest and only marginally significant differences in two out of twelve coefficients. This indicates that the parallel trends assumption required for causal interpretation of our difference-in-differences coefficients below is plausible.

16. By splitting the pre-period in half when studying pre-trends, we attempt to maximize estimate precision.

TABLE 3. Parallel Trends in Town Panel

	Post x Treatment		
Dependent variable	*Coefficient*	*Std. error*	*No. observations*
A. Number of accidents			
Total accidents, all vehicles	2.754	2.362	4,110
Motorcycle accidents	1.700	1.415	4,110
Serious motorcycle accidents	0.220	0.215	4,110
Minor motorcycle accidents	0.673	0.537	4,110
B. Shares (%)			
Helmet use in motorcycle accidents	–0.096	0.099	1,192
Serious moto over total accidents	0.044	0.159	1,385
Serious moto over moto accidents	0.122	0.161	1,192
Minor moto over moto accidents	–0.113	0.077	1,192
C. Characteristics of drivers and accidents			
Youth (< 25 years)	0.092*	0.053	1,372
Male	–0.028	0.129	1,383
Accident in urban area	0.005	0.050	1,382
Accident at night	–0.129*	0.077	1,385

*$p < 0.1$.

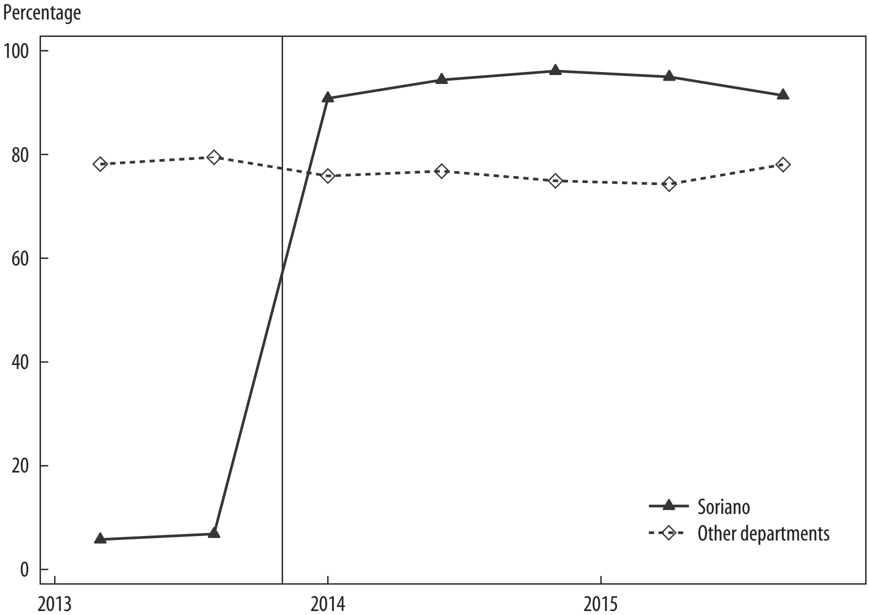
Notes: The table reports the coefficient identifying differences in dependent variable trends between treatment and comparison groups in 2013. Estimates were obtained using a town-month panel from January to October 2013. In panel B, for helmet use and serious and minor motorcycle accidents over total motorcycle accidents, the sample is restricted to town-month pairs with reported motorcycle accidents. For panel C and serious motorcycle accidents over total accidents in B, the sample is restricted to town-month pairs with reported road accidents. All regressions control for month and town fixed effects, with the exception of accidents in urban areas, where regressions control for month and department fixed effects. Standard errors are clustered at the town level.

Further evidence on the absence of substantial differences in the pre-trends of our outcomes is presented in figures 4 and 5, built using our individual accident data and averaging within groups. Figure 4 describes the evolution of average helmet use in Soriano and the rest of the country. In both figures, we report five-month averages of the corresponding outcome so that this is comparable to the exercise reported in table 2, and we average the high-frequency fluctuations in the outcome over time.[17] The overarching message from these figures is the same as for figures 2 and 3: before the policy, there were no sizable differences in trends between helmet use and serious accidents in the treatment and comparison groups, and this changed abruptly in 2014.

To obtain quantitative estimates of the effect of helmet use enforcement, we follow two different strategies. First, we use our localities panel to obtain difference-in-differences estimates of the effect of the change in enforcement

17. The total number of accidents in the comparison group is much larger than in Soriano, which results in a much smoother pattern at higher time frequencies.

FIGURE 4. Helmet Use in Soriano and the Rest of Uruguay



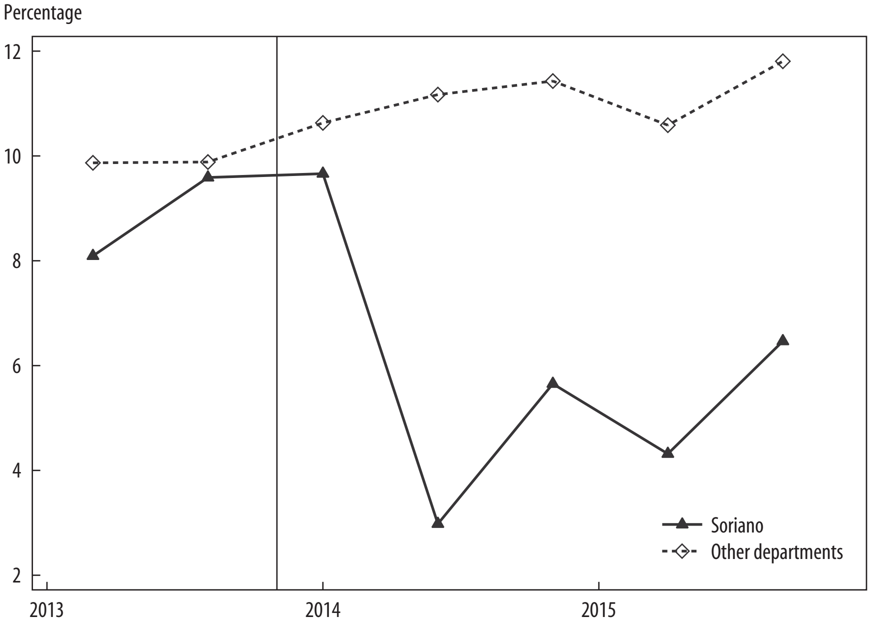
Source: Authors' calculations, based on data from UNASEV.
Notes: Helmet usage is measured as the percentage of all reported motorcycle accidents in which the driver was wearing a helmet. Vertical line corresponds to November 2013. Frequency: five-month averages.

on accident volumes and the types of accident occurring in different locations. For this purpose, we estimate a version of equation 1 in which the variable $Post_t$ is a dummy variable taking a value of one in the months after November 2013. Second, we use our data at the individual level to study the effect of enforcement on helmet use and accident severity. For this purpose, we restrict our sample to motorbike accidents and estimate the following:

$$Y_{it} = \alpha_j + \delta_t + \eta Post_t T_j + \varepsilon_{it}, \tag{2}$$

where i is an index for individuals involved in an accident, $Post_t$ takes a value of one in the months after November 2013, and T_j takes a value of one if the accident took place in the department of Soriano. When we use our accident-level data set, the outcome Y_{it} is either a dummy variable taking a value of one if the biker was wearing a helmet, a dummy variable taking a value of one

FIGURE 5. Serious Injuries in Soriano and the Rest of Uruguay



Source: Authors' calculations, based on data from UNASEV.
Notes: Serious injury or fatality is defined as a percentage of the number of motorcycle accidents in each city. Vertical line corresponds to November 2013. Frequency: five-month averages.

if the outcome from the accident was a serious injury, or a dummy variable taking a value of one if the outcome was a minor injury.

Finally, we can exploit the policy as a source of exogenous variation in helmet use to study the effect of helmet use on accident severity. To do so, we use the policy as an instrument for helmet use, so that equation 2 with a helmet dummy outcome is our first stage, and our second stage is given by

$$Severity_{it} = \alpha_j + \delta_t + \pi Helmet_{it} + \varepsilon_{it}. \tag{3}$$

The additional assumption in this particular exercise is the exclusion restriction: the change in policy affected accident severity only through its impact on helmet use. Several results in the next section indicate that this may be a reasonable assumption in our context.

TABLE 4. Number and Type of Accidents in All Locations

Estimation and variable	*(1)*	*(2)*	*(3)*
A. Accidents by vehicle			
Dependent variable	Total accidents	Moto accidents	Other vehicles
Post × Treatment	–0.034	–0.053	0.018
	(1.722)	(0.625)	(1.155)
No. observations	14,796	14,796	14,796
B. Accidents by cause			
Dependent variable	Collision	Falling	Other
Post × Treatment	–0.066	0.059	0.007
	(0.075)	(0.069)	(0.052)
No. observations	5,319	5,319	5,319

Notes: The estimates in panel A are obtained from a month-locality panel including locality fixed effects and year-month effects. In column 1, the dependent variable is the total number of people involved in traffic accidents in a locality-month pair; in column 2, the total number of people involved in motorcycle accidents; in column 3, the number of people involved in other vehicle accidents. The estimates in panel B are obtained from a month-locality panel including localities with at least one accident in a month-locality pair. The dependent variable is the share of motorcycle accidents arising from collisions, falling (for example, from the motorcycle), and other causes, respectively. All specifications include locality fixed effects and year-month effects. Standard errors clustered at the locality level are in parentheses.

Difference-in-Differences: Results

Difference-in-differences estimates for the effect of the change in enforcement on accident volumes for different vehicles are reported in panel A of table 4. Point estimates are small in absolute value in all columns, at less than 0.01 of a standard deviation of the dependent variable. They are also statistically insignificant at conventional levels. We interpret these findings as evidence that the enforcement of helmet use in Soriano had no impact on total accidents, motorbike accidents, or accidents involving other vehicles.

Results for accident types are reported in panel B of the table. In this case, we compute the share of all accidents corresponding to collisions, falling (for example, from a motorbike), or other causes. We again find no statistically significant effect of increased enforcement on the type of accident taking place.[18] These findings are important because they suggest that changes in perceived risks for motorcyclists, as a result of changes in enforcement, did not have substantial effects on risk taking or observable measures of driver behavior, as predicated by hypotheses of risk compensation by drivers.

18. The share of accidents by type is only defined for locality-month pairs featuring at least one accident. This implies that the sample used to produce the estimates in panel B of table 3 is heavily selected. However, the fact that there is no effect of increased enforcement on accident volumes implies that this sample selection should not have a substantial effect on our estimates.

TABLE 5. Difference-in-Differences Estimates for all Locations

Estimation method and variable	*(1)*	*(2)*
A. First stage		
Post × Treatment	0.902***	0.887***
	(0.030)	(0.037)
B. Reduced-form		
Post × Treatment	–0.047***	–0.049***
	(0.012)	(0.011)
C. 2SLS estimates (IV)		
Helmet D.	–0.052***	–0.055***
	(0.012)	(0.011)
No. observations	60,689	60,689
Vehicle	Motorbike	Motorbike
Month-year fixed effects	Yes	Yes
Town fixed effects	No	Yes

***$p < 0.01$.

Notes: In panel A, the dependent variable is a dummy taking a value of one if the victim of the accident was reportedly wearing a helmet at the time of the accident; in panels B and C, the dependent variable is a dummy taking a value of one if the accident victim experienced a serious or fatal injury. Panel C reports instrumental variable (IV) estimates of the effect of helmet use on serious accidents as discussed in the text. All estimates use the subsample of all registered motorcycle accidents. Standard errors clustered at the locality level are in parentheses.

We now turn to our individual-level data to obtain estimates of the effect of the change in enforcement on helmet use and accident severity. These are reported in table 5, where column 1 accounts for cross-sectional differences between treatment and comparison groups using a treatment dummy variable, and column 2 includes a full set of town dummy variables. In panel A, the coefficients show an increase of roughly 90 percent in helmet use as a result in the change in enforcement. This is in line with the results illustrated in figure 4, indicating that helmet use in Soriano went from close to zero to almost full compliance in a few months. Panel B provides reduced-form results for the effect of the enforcement of the mandatory helmet law on serious accidents. We find a negative and significant effect of –0.047, showing that the probability that a motorbike accident will result in a serious injury was reduced by approximately 4.7 percentage points as a result of the policy. This effect is large, as the baseline probability of having a serious or fatal injury for bikers is 11.3 percent in this sample.

Panel C of the table shows our instrumental variables (IV) estimates of the causal effect of helmet use. These roughly coincide with the ratio between the reduced-form coefficients in panel B and the first-stage estimates in panel A. The effect of interest is roughly 5 percent, indicating that helmet use reduces

TABLE 6. Minor Injuries and Helmet Use

Explanatory variable	*(1)*	*(2)*
Helmet use	0.037**	0.060***
	(0.018)	(0.014)
No. observations	54,213	54,213
Vehicle	Motorbike	Motorbike
Time fixed effects	Yes	Yes
Town fixed effects	No	Yes

$^{**}p < 0.05$; $^{***}p < 0.01$.

Notes: The dependent variable in all specifications is a dummy variable taking a value of one if the accident resulted in a minor injury and zero if the driver was unharmed. The estimation method is 2SLS, using the subsample of all registered motorcycle accidents. Standard errors clustered at the locality level are in parentheses.

the probability that a motorbike accident results in a serious or fatal injury by about 40 percent. This estimated effect is slightly *larger* than the difference in probability of serious injury obtained from the mean comparison in table 1. This suggests that helmet use is positively correlated with determinants of serious accident risk at the local level, such as local density and urbanization.

The reduction in the prevalence of serious injuries as a result of motorbike accidents can operate through either a change in the type of accident in which bikers are involved, or a change in accident severity conditional on accident type. Table 4 showed that the type of motorcycle accident does not change with the enforcement of helmet use. If changes in accident severity are driving the effect on serious injuries, we would expect a positive effect on minor injuries as a result of the change in enforcement. Accidents that would have resulted in a serious injury if a helmet was not used may result in a minor injury instead. To explore this, we reproduce the previous analyses using an indicator taking a value of one if an accident results in minor injuries and zero if the driver is unharmed as the dependent variable.[19] Results are reported in table 6. Instrumental variable estimates indicate that helmet use leads to a positive and significant effect on minor injuries, pointing to a transfer of serious to minor injuries as a result of the change in enforcement.

Based on the results reported in tables 5 and 6, we conclude that enforcement of the mandatory helmet use law led to a reduction in serious or fatal accidents and an increase in accidents resulting in minor injuries. We interpret this as a concomitant change in the relative probabilities of both types of

19. Including serious injuries among the zeroes does not change the qualitative results of the exercise.

accident. The fact that there are no discernible changes in the volume and type of accidents suggests that there are no other first-order behavioral responses to the law, at least in terms of driver behavior.[20] Therefore, we find that helmet use reduces accident severity and detect no evidence in support of the type of risk-compensating behavior associated with the Peltzman hypothesis.

Heterogeneous Effects

In this section, we study whether the policy change had different effects for different types of accident or victim. That is, we study whether the results from panel C in table 5 are heterogeneous across groups by reporting treatment effects for five subsamples, defined by age group, sex, being a driver (versus passenger), whether the accident occurred in an urban area, and whether it occurred at night. We estimate equation 3 for each group by splitting the sample according to each characteristic and running a two-stage least squares (IV) regression where helmet use is instrumented with a dummy variable for treated localities in the post-treatment period. Table 7 shows the results for eleven different subsamples.[21]

The first-stage results—not reported here, but available in the replication files—show that take-up is fairly uniform across subsamples, at around 0.9 for males and females, different age groups, accidents in urban or rural areas, and accidents during the night or during the day. The only subsample that has a lower take-up (0.6) is passengers (as opposed to drivers). In light of the findings in Grimm and Treibich (2016), these results indicate that the population induced to wear a helmet by enforcement of the corresponding law may differ substantially from the population of drivers who decide to wear a helmet spontaneously.

The broad picture of results from table 7 is that the benefits of helmet use on serious injuries are higher for the high-risk groups or accident types. When we split the sample by the age of the driver (columns 1 to 3) we find larger effects of helmet use on the young and old, and no effect at all on the middle-aged (between twenty-five and fifty-five years old). This may result from differences in risk attitudes and vulnerability by age, with young individuals being less risk averse (as shown in Dohmen and others, 2017) and relatively

20. Using a subsample of the UNASEV data set, we also explore the effect of the change in enforcement on the number of pedestrians involved in traffic accidents. Difference-in-differences estimates are negative, small, and statistically insignificant (results available on request).

21. Qualitative results are similar if, instead of splitting the sample, we use a less flexible model with an interaction of the treatment variable and each observable characteristic.

TABLE 7. Heterogeneous Treatment Effects

Explanatory variable	*(1) Youth*	*(2) Adults*	*(3) Seniors*	*(4) Males*	*(5) Females*	*(6) Drivers*
Helmet use	–0.081***	–0.007	–0.127***	–0.063***	–0.038***	–0.061***
	(0.001)	(0.018)	(0.014)	(0.015)	(0.010)	(0.009)
No. observations	24,544	25,996	6,770	41,991	18,535	51,437
	(7) Passengers	*(8) Urban*	*(9) Rural*	*(10) Night*	*(11) Day*	
Helmet use	–0.017	–0.058***	0.042	–0.071***	–0.046***	
	(0.040)	(0.010)	(0.040)	(0.013)	(0.017)	
No. observations	9,160	51,830	8,859	16,826	43,863	

***$p < 0.01$.

Notes: The table shows two-stage least squares estimates, with locality fixed effects. Helmet usage is instrumented with a dummy equal to one for accidents in the treated localities in the post-treatment period. The dependent variable in all specifications is a dummy taking a value of one if the accident victim experienced a serious or fatal injury. In columns 1 to 3, we split the sample by age group: youth: under twenty-five years old; adults: from twenty-five to forty-nine years old; and seniors: fifty years old or older. In columns 4 and 5, we split the sample by sex. In columns 8 and 9, we split the sample by the type of locality where the accident occurred. In columns 10 and 11, the night or day variable is set for each day, based on the time of sunset. All columns include a full set of time fixed effects. The sample comprises all registered motorcycle accidents. Standard errors clustered at the locality level are in parentheses.

older drivers being more physically vulnerable. We find larger coefficients (in absolute value) for males than females and for accidents occurring at night. Globally, we interpret these findings as suggesting that helmets are particularly important for subpopulations that are more at risk of injury.

Robustness Checks

In this section, we evaluate the robustness of the qualitative findings reported above by conducting five sets of complementary exercises: (1) we validate the inference methods above using spatial heteroskedastic and autocorrelation consistent (HAC) standard errors for our reduced-form estimates of the change in enforcement; (2) we provide two falsification tests—one using Cerro Largo instead of Soriano as the treatment group and the other focusing on car accidents—for our main results; (3) we obtain alternative estimates using a triple interaction model accounting for differences in accident rates across all vehicles; (4) we exclude either Cerro Largo or Montevideo from the comparison group; and finally (5) we use a synthetic control for Soriano (see also appendix B).[22]

22. Supplementary material for this paper is available online at http://economia.lacea.org/contents.htm.

Spatial HAC Standard Errors

Throughout most of the analysis above, our inference is carried out using standard errors clustered at the level of individual localities. This is motivated by the fact that it is likely that there are locality-level shocks to our dependent variables—accident volumes, helmet use, and accident outcomes. Yet the choice to cluster at the level of localities has two issues. First, our treatment varies at the department level, not at the locality level. Since Bertrand, Duflo, and Mullainathan (2004), much of the difference-in-differences literature obtains standard errors clustered at the level of treatment, but this was not feasible in our case because there are only nineteen departments in our sample.[23] Second, it is likely that our outcomes feature non-negligible spatial autocorrelation, so residuals in neighboring clusters will typically be correlated, violating the key assumption invoked to justify clustering at that level.

To address potential concerns with inference in our main tables, we report standard errors obtained using the spatial heteroskedasticity and autocorrelation consistent (HAC) robust standard errors proposed by Conley (1999), which are frequently used in much of the empirical literature in spatial economics. These standard errors are obtained by specifying a (typically uniform) spatial kernel and using these kernel weights to compute a variance-covariance matrix incorporating spatial dependence, analogous to an adjustment for heteroskedasticity and autocorrelation. Results for reduced-form difference-in-differences estimates on helmet use, the probability of an accident resulting in a serious injury, and the probability of an accident resulting in a minor injury are reported in table A4 in the online appendix. We use a spatial kernel 100 km in radius, so that the area of the uniform kernel is almost twice the size of the largest department in the country.[24] The main conclusions of our analysis are maintained with this inference method.

Falsification Tests

We can use our accidents data to build two suitable placebos in order to validate our methodology. First, we can use the department of Cerro Largo as a placebo

23. A growing literature proposes methods to conduct inference in the difference-in-differences setting when the number of clusters is small. However, these methods generally require having a large number of treated clusters, which is not the case in our paper (see MacKinnon and Webb, 2020).

24. The adjustment is carried out using the reg2hdfe spatial Stata command by Thiemo Fetzer (Fetzer, 2014), which is itself based on the previous implementation by Solomon Hsiang (Hsiang, 2010). We thank these authors for making these codes available.

to test whether there were changes in either helmet use or accident severity in this department coinciding with the introduction of the Misdemeanors Act in 2013. For this purpose, we reproduce the equivalent of our reduced-form estimates using this department as the treatment and all other departments—excluding Soriano—as the comparison group. Results for this exercise are presented in table A5 in the online appendix. As expected, we find no evidence of a significant effect of the interaction term on serious accidents. While the local governments of Cerro Largo and Soriano both refused to enforce helmet use by motorcyclists before late 2013, it is only in Soriano—which changed enforcement in that period—that we observe a substantial change in accident severity.

Second, we can use data on automobile accidents to study whether changes in the severity of these accidents responded to the change in policy in Soriano. We can only interpret this as a placebo if we assume that the change in helmet use does not affect the risks associated with car accidents. This assumption is perhaps reasonable given the results on accident volumes in table 4, although accident volumes might not sufficiently capture all of the possible changes in driver behavior or risks. The results indicate that the change in helmet use enforcement was not associated with changes in the severity of automobile accidents (see table A5).

Triple Differences Model

Our baseline estimates are obtained by focusing specifically on motorbike accidents. This is motivated by our interested in the effect of helmet use on the health outcomes of the motorcyclist involved in the accident itself. However, we can use a larger sample including all accidents to obtain similar estimates in a triple-interaction model. The advantage of this alternative specification is that it can help us account for potential time-varying confounders that differentially affect all accidents in the treatment and comparison groups, such as broader trends in road behavior or idiosyncratic changes in the intensity of all forms of road regulation. To account for overall shifts in accidents across vehicle types when estimating our effect of interest, we estimate the following equation:

$$
\begin{aligned}
(4) \qquad Y_{it} = \alpha_j + \delta_t + \beta Post_t T_i Moto_{it} + \gamma_1 Post_t Moto_{it} + \gamma_2 Moto_{it} T_i \\
+ \gamma_3 T_i Post_t + \varepsilon_{it}.
\end{aligned}
$$

where T_i and $Post_t$ are defined as above and Y_{it} is either a dummy variable that takes a value of one if the motorcyclist in accident *i* was wearing a helmet or a dummy variable that takes a value of one if the motorcyclist suffered a minor accident. The variable $Moto_{it}$ takes a value of one if the victim involved in the accident is a motorcyclist. As in our baseline difference-in-differences specification, we also use variation in enforcement (captured by the triple interaction term) as an instrument for helmet use to obtain an estimate of the effect of helmet use on accident severity. The innovation relative to the specification in equation 2 comes in the form of the interaction term T_iPost_t, which accounts for changes over time in accident severity of all vehicles between treatment and comparison groups.

Estimates for the coefficient on the triple interaction term for the first-stage, reduced-form, and IV specifications are reported in table A6 in the online appendix. Results are broadly consistent with those reported in table 5 using motorcyclists only. We interpret this as evidence that our baseline results are not driven by factors unrelated to helmet use enforcement affecting all vehicle accidents.

Alternative Comparison Groups

In this section, we test whether our results are robust to specific choices regarding the composition of the comparison group. Our baseline estimates use motorbike accidents in all recorded locations. However, certain locations may be ill-suited to act as controls. Cerro Largo, for example, is different from other departments because it did not enforce helmet use throughout the whole period. More important, the capital city of Montevideo is the largest urban area in the country, is characterized by a relatively more modest use of motorcycles, and has a high-density environment that is quite distinct from other localities in the country (see table 2).

Table A7 in the online appendix presents IV estimates of the effect of helmet use on the probability of minor and serious injuries after excluding accidents taking place in Cerro Largo and Montevideo. Comparing these estimates with those reported in tables 4 and 5 indicates that these sample restrictions have little impact on our findings.

We can alternatively restrict our sample to accidents taking place in Mercedes and Melo only, so as to provide quantitative estimates of the effects illustrated in figures 2 and 3. Results for this exercise are qualitatively and quantitatively in line with those reported earlier (see table A8 in the online appendix).

Synthetic Control

The difference-in-differences estimates reported in the previous sections result from comparing changes in an outcome (for example, serious accident rate) between locations in Soriano and the rest of the country. These control groups are natural choices, but they are also arbitrary. We can use the data-driven synthetic control method—as described in Abadie, Diamond, and Hainmueller (2010)—to select a suitable control group to use to estimate the difference in the rate of serious injuries induced by the policy.

Online appendix B discusses the implementation and results of applying this method using aggregate department-level data. Our results are qualitatively in line with our findings in the difference-in-differences analysis reported earlier. Soriano experienced a sustained reduction in accumulated serious motorbike accidents per capita after the fourth quarter of 2013. This reflects a change in accident severity and not accident volumes, which remained relatively stable throughout the period.[25]

Discussion

We can use our estimates and additional information on health and administrative costs to outline a cost-benefit analysis of helmet use laws for Uruguay. The main benefits of the policy arise from the reduction in serious injuries and fatalities from motorcycle accidents. The main costs relate to the administrative costs of enforcement paid by the relevant agencies and the nuisance costs of wearing a helmet for motorcyclists. The latter is particularly hard to estimate, but we can calculate the magnitude of these costs that would reverse the change in benefits.[26] The outcome of the cost-benefit analysis can then be obtained relative to this benchmark.

The health benefits of the change in enforcement derive from a reduction in the volume of serious accidents and deaths. Paolillo and others (2016) document that roughly one and a half out of ten serious traffic accidents lead to a fatality. They also estimate the average intensive care hospitalization costs for serious traffic accidents in Uruguay to be USD 7,437. A conservative estimate

25. See online appendix B for details.

26. Standard revealed-preference valuation tools, such as the opportunity cost or compensating differential methods, cannot be applied to measure nuisance costs because there are no other markets compensating for these costs or pricing a similar bad.

for the value of a statistical life is USD 2,346,000.[27] We obtain health benefits by multiplying these figures times an estimate of the absolute reduction in serious injuries. The coefficient on the reduced-form effect of the policy on serious accidents in column 2 of table 5 is 4.86 percent. The average number of yearly motorcycle accidents in Soriano is 610. Hence the policy leads to a reduction of roughly twenty-nine serious or fatal accidents per year. Using this number, we can compute the estimated health benefits from the policy as $29 \times 0.15 \times 2{,}346{,}000 + 29 \times 0.85 \times 7{,}437$. This yields a figure of USD 10,389,727 per year in benefits arising from reduced hospitalization costs and deaths alone. Assuming a 5 percent discount rate and a thirty-year time horizon (as in Dee, 2009), the present value of health benefits would be on the order of USD 160 million. This corresponds to USD 6,053 per motorcyclist.

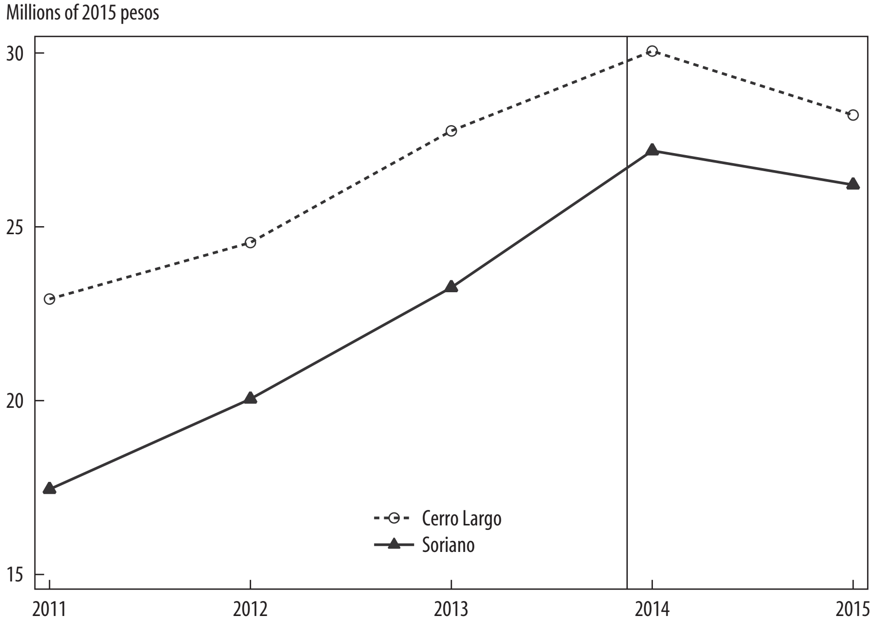
Other health effects, such as the psychological costs and permanent disability resulting from serious accidents or the reduced work hours for hospitalized patients, are likely to be substantial. Therefore, we consider these figures to be an underestimate of total health benefits.[28]

On the cost side, public enforcement of the helmet law requires the use of traffic inspectors to detect and sanction violators. How much of Soriano's public resources were devoted to these tasks? Figure 6 reports the personnel expenses of the Soriano and Cerro Largo Transit Departments. The parallel trends observed before the enforcement of the law do not change afterward. In other words, Soriano achieved an abrupt increase in compliance with the helmet law after 2013 without an escalation in personnel costs. Consulted officials at the Soriano transit authority stated that enforcement of the law did not involve the deployment of additional human resources. Inspectors were already deployed within the city of Mercedes to enforce other transit rules (such as speed limits and traffic lights), and after the law was enforced,

27. In the literature, there is considerable uncertainty about the value of life, depending on the method used, the age of the victim, and the country where it is estimated. According to the U.S. EPA (2014), a recommended default central value of a statistical life (VSL) is around USD 8.7 million (in 2014 dollars). The U.S. Department of Transportation (2013) indicates that, on the basis of the best available evidence, the VLS that should be used for calculating the benefits of preventing traffic fatalities is USD 9.1 million (in 2012 dollars). Considering that Uruguayan GDP per capita is 27 percent of U.S. GDP per capita, we employ a conservative value of USD 2,346,000 for our estimates.

28. As discussed earlier, the reduction in serious and fatal injuries comes at the expense of an increase in minor injuries. Minor injuries will impose costs of their own, although by definition they will not require hospitalization. These unaccounted costs are arguably higher for serious accidents, so our estimate of the net health benefits would still be a lower bound of total health costs, even after accounting for the increased number of minor injuries.

FIGURE 6. Personnel Expense of the Municipal Transit Department



Source: Observatorio Territorio Uruguay (OPP).
Note: Vertical line corresponds to November 2013.

the same inspectors just added another complementary task—enforcement of the helmet law—to their daily activity. Information campaigns on helmet use were included in traffic safety campaigns already in place before the policy change. Hence it is not surprising that we do not identify a significant administrative cost of enforcement in this case.

With regard to the nuisance costs of helmet use, there were 26,435 registered motorcycles in Soriano in 2013. The nuisance costs for registered motorcycles resulting from the policy should be proportional to this figure, scaled by the change in helmet use, which is 89 percent (see table 5). Our health benefits estimate is USD 10,389,727 per year, so the policy would have a positive net benefit for yearly nuisance costs per registered motorcycle under USD 442. Because our estimate of health benefits is probably downward biased, this is a lower bound for break-even nuisance costs per motorcyclist.

In light of this discussion, low levels of helmet use in the absence of appropriate enforcement in 2013 can be explained on three grounds: large nuisance costs, moral hazard, or biased risk perception. First, if the nuisance costs of wearing a helmet—plus the pecuniary costs of owning one—are well above USD 442 a year, then the laissez-faire outcome is that rational motorcyclists will choose not to wear a helmet. Second, motorcyclists may not internalize the full cost of serious injuries because of the pervasiveness of health and disability insurance. If this is the case, even if the costs of helmet use are below USD 442 per year, it may still be privately optimal for drivers not to use a helmet. Finally, it is not obvious that motorcyclists have an accurate perception of the risks of driving without a helmet. The same outcome of low helmet use without enforcement would be observed if motorcyclists' subjective probabilities of serious accidents are lower than actual probabilities.

Conclusion

Mandatory helmet use laws for motorcyclists are a feature of transit regulation in many jurisdictions. They are not universal, however, and enforcement is often extremely poor, particularly in low- and middle-income countries. This paper shows that changes in enforcement can lead to a substantial alleviation of the deleterious health consequences of motorcycle accidents. Our difference-in-differences estimates indicate that changes in the enforcement of helmet use laws in Uruguay led to a substantial reduction of roughly five percentage points in the rate of serious or fatal injuries. Since the national base rate stands at roughly 11 percent for this period, this effect is sizable. The reduction in serious accidents takes place at the expense of an increase in minor injuries, pointing squarely to a net reduction in accident severity. Accident numbers and the type of accident taking place—for both motorcycles and other vehicles—do not appear to be affected by the change in policy. This further alleviates concerns that behavioral responses to helmet use in the form of risk-compensating actions—such as increased driving speeds or more reckless conduct by motorcyclists—counter the direct effect of using a helmet to prevent head trauma.

Combining our reduced-form estimates of changes in accident severity with hospitalization costs and the value of statistical life, we calculate an approximate measure of the health benefits resulting from the change in enforcement. Because the direct enforcement costs for the involved traffic control agencies were largely unaffected by the policy, the main costs of increased helmet use are associated with the potential nuisance for riders. However, the nuisance costs would have to be substantial to offset the policy's health benefits.

References

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller. 2010. "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program." *Journal of the American Statistical Association* 105 (490): 493–505.

Aney, Madhav S., and Christine Ho. 2019. "Deadlier Road Accidents? Traffic Safety Regulations and Heterogeneous Motorists' Behaviour." *Regional Science and Urban Economics* 77 (July): 155–71.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan. 2004. "How Much Should We Trust Differences-in-Differences Estimates?" *Quarterly Journal of Economics* 119 (1): 249–75.

Carpenter, Christopher S., and Mark Stehr. 2008. "The Effects of Mandatory Seatbelt Laws on Seatbelt Use, Motor Vehicle Fatalities, and Crash-Related Injuries among Youths." *Journal of Health Economics* 27 (3): 642–62.

———. 2011. "Intended and Unintended Consequences of Youth Bicycle Helmet Laws." *Journal of Law and Economics* 54 (2): 305–24.

Chong, Alberto, and Pascual Restrepo. 2017. "Regulatory Protective Measures and Risky Behavior: Evidence from Ice Hockey." *Journal of Public Economics* 151 (July): 1–11.

Cohen, Alma, and Liran Einav. 2003. "The Effects of Mandatory Seat Belt Laws on Driving Behavior and Traffic Fatalities." *Review of Economics and Statistics* 85 (4): 828–43.

Conley, Timothy G. 1999. "GMM Estimation with Cross Sectional Dependence." *Journal of Econometrics* 92 (1): 1–45.

Dee, Thomas S. 2009. "Motorcycle Helmets and Traffic Safety." *Journal of Health Economics* 28 (2): 398–412.

Dickert-Conlin, Stacy, Todd Elder, and Brian Moore. 2011. "Donorcycles: Motorcycle Helmet Laws and the Supply of Organ Donors." *Journal of Law and Economics* 54 (4): 907–35.

Dohmen, Thomas, Armin Falk, Bart H. H. Golsteyn and others. 2017. "Risk Attitudes across the Life Course." *Economic Journal* 127 (605): 95–116.

Evans, Leonard, 1986. "Double-Pair Comparison: A New Method to Determine How Occupant Characteristics Affect Fatality Risk in Traffic Crashes." *Accident Analysis and Prevention* 18 (3): 217–27.

Fetzer, Thiemo. 2014. "Can Workfare Programs Moderate Violence? Evidence from India." Economic Organisation and Public Policy Discussion Paper 053. London: London School of Economics, Suntory and Toyota International Centres for Economics and Related Disciplines (STICERD).

Grimm, Michael, and Carole Treibich. 2016. "Why Do Some Motorbike Riders Wear a Helmet and Others Don't? Evidence from Delhi, India." *Transportation Research Part A: Policy and Practice* 88 (June): 318–36.

Hansen, Benjamin. 2015. "Punishment and Deterrence: Evidence from Drunk Driving." *American Economic Review* 105 (4): 1581–617.

Houston, David J., and Lilliard E. Richardson. 2008. "Motorcyclist Fatality Rates and Mandatory Helmet-Use Laws." *Accident Analysis and Prevention* 40 (1): 200–08.

Hsiang, Solomon M. 2010. "Temperatures and Cyclones Strongly Associated with Economic Production in the Caribbean and Central America." *Proceedings of the National Academy of Sciences* 107 (35): 15367–72.

Lee, Jonathan M. 2018. "Mandatory Helmet Legislation as a Policy Tool for Reducing Motorcycle Fatalities: Pinpointing the Efficacy of Universal Helmet Laws." *Accident Analysis and Prevention* 111 (February): 173–83.

Loeb, Peter D. 1995. "The Effectiveness of Seat-Belt Legislation in Reducing Injury Rates in Texas." *American Economic Review* 85 (2): 81–84.

MacKinnon, James G., and Matthew D. Webb. 2020. "Randomization Inference for Difference-in-Differences with Few Treated Clusters." *Journal of Econometrics* 218 (2): 435–50.

Olsen, Cody S., Andrea M. Thomas, Michael Singleton, and others. 2016. "Motorcycle Helmet Effectiveness in Reducing Head, Face, and Brain Injuries by State and Helmet Law." *Injury Epidemiology* 3 (1): 23–28.

Paolillo, Elbio, Alberto Scasso, Frank Torres, and others. 2016. "Siniestros de tránsito, los grupos relacionados por el diagnóstico y los costos hospitalarios." *Revista médica Uruguaya* 32 (1): 41–50.

Peltzman, Sam. 1975. "The Effects of Automobile Safety Regulation." *Journal of Political Economy* 83 (4): 677–725.

Peng, Yinan, Namita Vaidya, Ramona Finnie, and others. 2017. "Universal Motorcycle Helmet Laws to Reduce Injuries: A Community Guide Systematic Review." *American Journal of Preventive Medicine* 52 (6): 820–32.

Pope, Adam T., and Robert D. Tollison. 2010. "Rubbin' Is Racin': Evidence of the Peltzman Effect from NASCAR." *Public Choice* 142 (3–4): 507–13.

U.S. EPA (Environmental Protection Agency). 2014. *Guidelines for Preparing Economic Analyses*. Washington, DC: EPA.

U.S. Department of Transportation. 2013. *Guidance on Treatment of the Economic Value of a Statistical Life.* Washington, DC: DOT.

Van Benthem, Arthur. 2015. "What Is the Optimal Speed Limit on Freeways?" *Journal of Public Economics* 124 (April): 44–62.

WHO (World Health Organization). 2006. *Helmets: A Road Safety Manual for Decision-Makers and Practitioners*. Geneva: WHO.

———. 2018. *Global Status Report on Road Safety 2018*. Geneva: WHO.

Latin American and Caribbean Economic Association
www.lacea.org

Brookings Institution Press
Washington, D.C.
www.brookings.edu

Cover design by Rogue Element