



Munich Personal RePEc Archive

# Measuring the size of the shadow economy using a dynamic general equilibrium model with trends: a new dataset

Chung, Federico and Purkey, Liam and Solis-Garcia, Mario

Macalester College, Macalester College, Macalester College

28 February 2020

Online at <https://mpra.ub.uni-muenchen.de/98963/>  
MPRA Paper No. 98963, posted 08 Mar 2020 08:29 UTC

# Measuring the size of the shadow economy using a dynamic general equilibrium model with trends: a new dataset

Federico Chung\*  
Macalester College

Liam Purkey†  
Macalester College

Mario Solis-Garcia‡  
Macalester College

February 28, 2020

## Abstract

We provide estimates of the size and dollar value of shadow economy for a set of countries between 1950 and 2015, following the methodology of [Solis-Garcia and Xie \(2018\)](#).

**JEL codes:** E26, E32, O17.

**Keywords:** informal sector, business cycles, DSGE models.

---

\* E-mail: [fchung@macalester.edu](mailto:fchung@macalester.edu).

† E-mail: [lpurkey@macalester.edu](mailto:lpurkey@macalester.edu).

‡ Corresponding author. E-mail: [msolisga@macalester.edu](mailto:msolisga@macalester.edu).

# 1 Introduction

This paper builds on the methodology of [Solis-Garcia and Xie \(2018\)](#) and derives estimates of the size and monetary value of the shadow economy for a panel of countries between 1950 and 2015. In a nutshell, the methodology uses the restrictions imposed by a full-fledged dynamic general equilibrium (GE) model to measure the size and dynamic behavior of the shadow economy. While earlier GE-based contributions have attempted to do this, they all omit one key factor, which is to consider the trends that are observed in the formal (measured) economy.<sup>1,2</sup>

To see why this matters, consider the works of [Ihrig and Moe \(2004\)](#) or [Elgin and Öztunalı \(2012\)](#); both of these are highly-cited and GE-based works. The former considers stationary economies only, which forego all the information contained in the real-world trends. The latter imposes dynamic relationships between the model’s trend growth rates that are not necessarily equilibrium rates. As shown by [Solis-Garcia and Xie](#), incorporating the trends of the model into the analysis provides a very different picture of the underground economy.

The main contribution of the paper is a dataset that offers time series for (a) formal-sector real GDP (hereafter RGDP), (b) shadow-sector RGDP, (c) total (formal plus shadow) RGDP, and (d) the size of the shadow economy as a fraction of the formal economy. The dataset includes 60 countries, yet data availability forces us to have staggered starting dates: 30 countries have data for the entire sample (1950–2015), 15 countries begin in 1970, and the remaining 15 do so in 1990.

The rest of the paper proceeds as follows. [Section 2](#) briefly reviews the methodology used to generate the values in the paper and [Section 3](#) presents some data preliminaries. Finally, [Section 4](#) presents some aggregated results and [Section 5](#) concludes. [Appendix A](#) contains the country

---

<sup>1</sup> Other papers use econometric techniques to infer the size of the shadow economy (e.g., [Medina and Schneider 2017](#) or [Schneider et al. 2010](#)); as documented in [Solis-Garcia and Xie \(2018\)](#), this methodology exhibits some issues when looking at the dynamics of the shadow economy.

<sup>2</sup> Simply put, the cross-equation restrictions that arise from GE effects generate strong implications about how the shadow economy can evolve over time. We exploit these implications in the paper.

classifications for all the Figures in Section 4, while Appendix B presents a table with summary statistics of the size of the shadow economy for all countries in the sample. All the data discussed in the paper can be found in XLSX format at the corresponding author’s website.<sup>3</sup>

## 2 A brief review of the methodology

This is an abridged description of the methodology we use to derive the results. For additional details, see Solis-Garcia and Xie (2018).

### 2.1 Economic environment

We consider a deterministic environment consisting of a representative household-producer and a government; in what follows, uppercase letters denote trending variables while lowercase letters denote stationary variables.

**Household-producer** The household-producer chooses sequences of consumption  $C_t$ , hours worked  $N_t$ , and investment  $X_t$  to maximize the discounted flow of period utilities, namely

$$\sum_{t=0}^{\infty} \beta^t \left( \log C_t - \frac{\Gamma_{Ht} N_t^{1+\chi}}{1+\chi} \right).$$

In the expression above,  $\beta \in (0, 1)$  is the discount factor,  $\chi \geq 0$  represents the inverse of the Frisch elasticity of labor supply, and  $\Gamma_{Ht}$  is a permanent shock that affects the household’s labor supply. The household-producer’s maximization is subject three constraints. First, a law of motion for capital

$$K_{t+1} = (1 - \delta)K_t + X_t, \tag{2.1}$$

---

<sup>3</sup> See <http://www.maclester.edu/~msolisga>.

where  $\delta \in (0, 1)$  is the depreciation rate and  $K_t$  denotes the stock of physical capital. Second, a time constraint of the form

$$N_t = N_{Ft} + N_{St}, \quad (2.2)$$

where  $N_{Ft}$  denotes hours worked in the formal sector and  $N_{St}$  does so for the shadow sector. Third, a budget constraint

$$C_t + \Gamma_{At}X_t = (1 - \tau_t)K_t^\alpha(\Gamma_{Ft}N_{Ft})^{1-\alpha} + (\Gamma_{St}N_{St})^\eta, \quad (2.3)$$

where  $\alpha \in (0, 1)$  is the capital income share in formal output,  $\eta > 0$  is the labor share in shadow output,  $\Gamma_{At}$  is a permanent shock to the production of investment goods, and  $\tau_t \in (0, 1)$  is a tax on formal sector output. The last term in (2.3) represents shadow sector output; importantly, we assume that the government cannot tax output produced underground.<sup>4</sup> Formal and shadow production technologies are subject to the permanent productivity shocks  $\Gamma_{Ft}$  and  $\Gamma_{St}$ .

**Government sector** The literature on the shadow economy (e.g., [Ihrig and Moe 2004](#)) has empirically shown that tax rates are negatively related to shadow sector size; we incorporate this feature of the data by including a government sector. In particular, we assume that the government uses tax revenue to pay for wasteful expenditure  $G_t$  and that it complies with the period-by-period budget constraint

$$G_t = \tau_t K_t^\alpha (\Gamma_{Ft} N_{Ft})^{1-\alpha}. \quad (2.4)$$

**Exogenous variables** The set of (permanent) exogenous variables is given by  $\Gamma_{Ht}$ ,  $\Gamma_{At}$ , and  $\Gamma_{Ft}$ ; [Proposition 2.1](#) below shows that  $\Gamma_{St}$  is an endogenous variable. Letting  $g_{it}$  denote the gross

---

<sup>4</sup> See [Solis-Garcia and Xie \(2018\)](#) for a discussion on why adding an audit probability to the shadow sector production doesn't change the results that follow.

growth rate of variable  $i = H, A, F, S$ , it's obvious that

$$g_{it} = \frac{\Gamma_{it}}{\Gamma_{i,t-1}},$$

and

$$\Gamma_{it} = \prod_{s=1}^t g_{is}. \quad (2.5)$$

## 2.2 Equilibrium

The equilibrium conditions of the economy are as follows. First, by the household-producer constraints described above:

$$C_t + \Gamma_{At}X_t + G_t = K_t^\alpha(\Gamma_{Ft}N_{Ft})^{1-\alpha} + (\Gamma_{St}N_{St})^\eta \quad (2.6)$$

$$K_{t+1} = (1 - \delta)K_t + X_t \quad (2.7)$$

$$N_t = N_{Ft} + N_{St}. \quad (2.8)$$

Second, we add an intertemporal condition

$$\Gamma_{At}C_t^{-1} = \alpha\beta C_{t+1}^{-1}(1 - \tau_{t+1})K_{t+1}^{\alpha-1}(\Gamma_{F,t+1}N_{F,t+1})^{1-\alpha} + \beta(1 - \delta)\Gamma_{A,t+1}C_{t+1}^{-1}, \quad (2.9)$$

two intratemporal conditions—for formal and shadow labor, respectively

$$\phi\Gamma_{Ht}N_t^X = (1 - \alpha)C_t^{-1}(1 - \tau_t)K_t^\alpha\Gamma_{Ft}^{1-\alpha}N_{Ft}^{-\alpha} \quad (2.10)$$

$$\eta\Gamma_{St}^\eta N_{St}^{\eta-1} = (1 - \alpha)(1 - \tau_t)K_t^\alpha\Gamma_{Ft}^{1-\alpha}N_{Ft}^{-\alpha}, \quad (2.11)$$

and the government budget constraint

$$G_t = \tau_t K_t^\alpha (\Gamma_{Ft} N_{Ft})^{1-\alpha}. \quad (2.12)$$

Third, we add expressions for formal ( $Y_{Ft}$ ), shadow ( $Y_{St}$ ), and total ( $Y_t$ ) output, namely

$$Y_{Ft} = K_t^\alpha (\Gamma_{Ft} N_{Ft})^{1-\alpha} \quad (2.13)$$

$$Y_{St} = (\Gamma_{St} N_{St})^\eta \quad (2.14)$$

$$Y_t = Y_{Ft} + Y_{St}. \quad (2.15)$$

These expressions provide a clean mapping between theory and data. Finally, we also derive an expression for the decentralized price of investment goods, given by

$$P_{Xt} = \Gamma_{At}. \quad (2.16)$$

### 2.3 The role of trend growth

**Theory** There are four model trends in the economy: the household's choice of hours worked ( $\Gamma_t^H$ ), the production of investment goods ( $\Gamma_t^A$ ), and formal and shadow technology productivity ( $\Gamma_t^F$  and  $\Gamma_t^S$ ). The main object of interest is the sequence  $\{\Gamma_t^S\}$ ; in order to derive a measure of the shadow economy, we first build a map between (equilibrium) observed and model growth rates, as detailed in [Proposition 2.1](#) below (all proofs can be found in [Solis-Garcia and Xie 2018](#)):

**Proposition 2.1.** *The equilibrium growth rates of the capital stock,  $g_K$ ; (formal and shadow) hours worked,  $g_N$ ; (formal, shadow, and total) output,  $g_Y$ ; and the permanent shock to the shadow*

production function,  $g_S$ , are given by

$$g_K = g_H^{-1/(1+\chi)} g_A^{-1/(1-\alpha)} g_F \quad (2.17)$$

$$g_N = g_H^{-1/(1+\chi)} \quad (2.18)$$

$$g_Y = g_H^{-1/(1+\chi)} g_A^{-\alpha/(1-\alpha)} g_F \quad (2.19)$$

$$g_S = g_H^{-(1+\eta)/[(1+\chi)\eta]} g_A^{-\alpha/[(1-\alpha)\eta]} g_F^{1/\eta}. \quad (2.20)$$

Equation (2.20) is key, as it links the evolution of the shadow sector to the exogenous growth rates of the model  $\{g_H, g_A, g_F\}$  along the equilibrium path. We now show how to calculate these growth rates from real-world data.

**Data** Let  $\{\hat{g}_K, \hat{g}_{N_F}, \hat{g}_{Y_F}\}$  denote the observed long-run averages of the growth rates of physical capital, formal hours worked, and formal output. The relation between theory and real-world rates follows

**Proposition 2.2.** *The map between the exogenous growth rates  $\{g_H, g_A, g_F\}$  and the observed growth rates  $\{\hat{g}_K, \hat{g}_{N_F}, \hat{g}_{Y_F}\}$  is given by:*

$$g_H = \hat{g}_{N_F}^{-(1+\chi)} \quad (2.21)$$

$$g_A = \hat{g}_{Y_F} \hat{g}_K^{-1} \quad (2.22)$$

$$g_F = \left( \frac{\hat{g}_{Y_F}}{\hat{g}_K^\alpha \hat{g}_{N_F}^{1-\alpha}} \right)^{1/(1-\alpha)}. \quad (2.23)$$



## 2.4 Parametrization and solution method

We set  $\alpha = 1/3$  and  $\chi = 1$ ,<sup>5</sup> we also need the following

**Assumption 2.3.** The observed (real-world) value of RGDP corresponds to formal output  $Y_{Ft}$ .

To obtain the value of the shadow sector labor input parameter  $\eta$ , we first need to pin down the value of the shadow-formal output ratio for some base year  $t_0$ ; call this value  $Y_{[S/F],t_0}$ . By construction,

$$Y_{[S/F],t_0} = \frac{Y_{S,t_0}}{Y_{F,t_0}} \quad (2.24)$$

and hence, using real-world data for formal output, shadow output at  $t_0$  is given by

$$Y_{S,t_0} = Y_{[S/F],t_0} Y_{F,t_0}. \quad (2.25)$$

Take (2.24) once more, but now substitute the definition of  $Y_{S,t_0}$ , equation (2.14):

$$Y_{[S/F],t_0} = \frac{(\Gamma_{S,t_0} N_{S,t_0})^\eta}{Y_{F,t_0}}.$$

From this expression we can easily solve for  $N_{S,t_0}$ , which equals

$$N_{S,t_0} = \frac{(Y_{[S/F],t_0} Y_{F,t_0})^{1/\eta}}{\Gamma_{S,t_0}}. \quad (2.26)$$

To wrap up, solve for  $\eta$  from the intratemporal condition (2.11):

$$\eta = \frac{(1 - \alpha)(1 - \tau_{t_0})Y_{F,t_0}N_{S,t_0}}{N_{F,t_0}Y_{S,t_0}}.$$

---

<sup>5</sup> The technical appendix to Solis-Garcia and Xie (2018) contains a sensitivity analysis over the values of  $\chi$ ; it also shows that dropping the assumption of logarithmic utility over consumption doesn't result in a major quantitative change.

and substitute from (2.25) and (2.26) to get<sup>6</sup>

$$\eta = \frac{(1 - \alpha)(1 - \tau_{t_0})Y_{F,t_0}^{1/\eta}Y_{[S/F],t_0}^{(1-\eta)/\eta}}{N_{F,t_0}\Gamma_{S,t_0}}. \quad (2.27)$$

Expression (2.27) is a nonlinear equation in  $\eta$ ; note that all the variables in the right-hand side of the equation—other than  $\eta$ —are known. We use a fixed point procedure to find the value of  $\eta$  such that (2.27) is satisfied (the fixed point algorithm is detailed in Solis-Garcia and Xie 2018).

### 3 Data preliminaries

#### 3.1 Sources

From the Penn World Table 9.0 (see Feenstra et al. 2015) we obtain the following variables (where appropriate, mnemonics are indicated in parenthesis):

1. Real GDP at constant 2011 national prices (rgdpna).
2. Capital stock at constant 2011 national prices (rkna).
3. Share of government consumption at current PPPs (cshg).
4. Price level of capital formation (p1i).

From The Conference Board,<sup>7</sup> we obtain the following variables:

5. Midyear population.
6. Total annual hours worked.

From the work of Medina and Schneider (2017), we get

<sup>6</sup> From (2.4) and period-by-period government budget balance we get that  $\tau_{t_0} = G_{t_0}/[K_{t_0}^\alpha(\Gamma_{F,t_0}N_{F,t_0})^{1-\alpha}] = G_{t_0}/Y_{F,t_0}$ ; the rightmost term can be backed out from real-world data.

<sup>7</sup> See The Conference Board *Total Economy Database*<sup>TM</sup>, May 2016, <http://www.conference-board.org/data/economydatabase/>.

7. Shadow-to-formal output ratio, with base year 1991.<sup>8</sup>

Using the items above, our model variables are obtained as follows:

8. Formal GDP per capita: 1/5.

9. Capital stock per capita: 2/5.

10. Price of investment: 4.

11. Hours worked per capita: 6/5.

12. Shadow-to-formal output ratio: 7.

### 3.2 Aggregation

The series presented in [Section 4](#) below are weighted averages of the shadow-to-formal output ratio, using country population as a weight. In each case, data from multiple countries are averaged according to criteria such as region, income level, and so on.

Each weighted series is constructed as follows. First, we partition the set of countries into  $M$  groups, denoted by  $\{G^1, \dots, G^M\}$ , with typical element  $G^m$ ; each group  $G^m$  has  $N^m$  elements, where we don't require that  $N^m = N^{m'}$  for  $m \neq m'$ . Second, let  $p_{jt}^m$  be the population of country  $j$  in group  $G^m$  at period  $t$ . Similarly, let  $w_{jt}^m$  be the weighing factor of country  $j$  at period  $t$ , defined as

$$w_{jt}^m = \frac{p_{jt}^m}{\sum_j p_{jt}^m}, \quad \text{for } j \in G^m.$$

(Note how each weighing factor is relative to a particular group  $G^m$ .) Finally, let  $Y_{[S/F],jt}^m$  be the shadow-to-formal output ratio of country  $j$  in group  $G^m$  at period  $t$ . Then the weighted shadow-

---

<sup>8</sup> See [Solis-Garcia and Xie \(2018\)](#) for an explanation on why we choose to use this source.

to-formal output ratio of group  $G^m$  in period  $t$  is

$$Y_{[S/F],t}^m = \sum_{j \in G^m} w_{jt}^m Y_{[S/F],jt}^m, \quad \text{for } j \in G^m.$$

## 4 Results

We now present time series for the size of the shadow economy, aggregated across four main dimensions: region, income level, trade block, and trade openness.<sup>9</sup>

### 4.1 By region

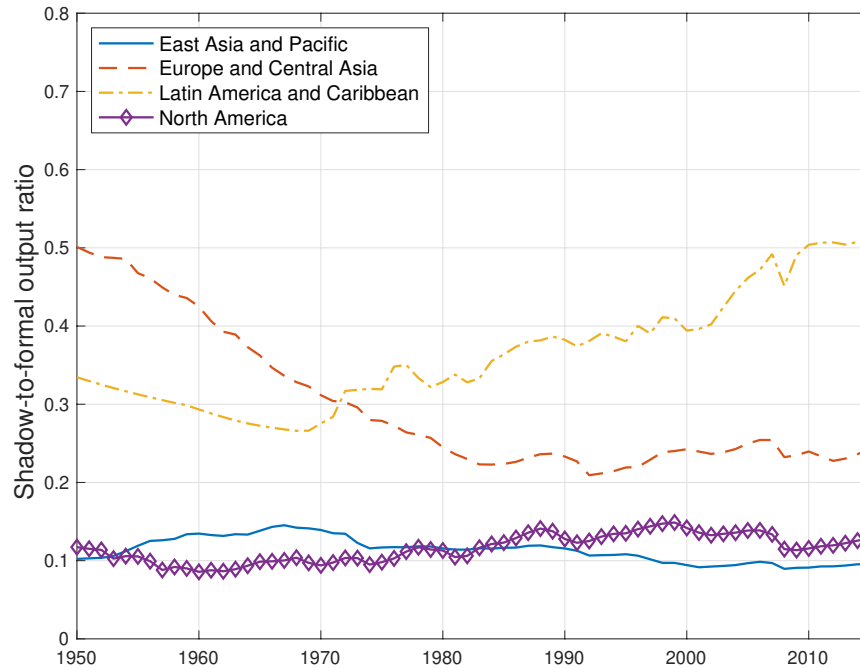
A limited sample of countries have data from 1950 onwards; consequently, some trends differ markedly once more countries are included in later time series.<sup>10</sup> The first point of interest in [Figure 1](#) is the marked decline of the shadow economy in Europe and Central Asia, of which the 1950 sample primarily consisted of western European countries, over the observed period. The decline continues up to 1980, after when it stabilizes around 20% of formal RGDP. Conversely, the size of the shadow economy in Latin America and the Caribbean has been on an upward trend since 1970—and has been consistently larger than all other regions during this period. The shadow economy in East Asia and Pacific and North America has remained small, hovering around 10% of formal RGDP.

The full sample from 1990 onward confirms some trends and reveals others that were not apparent in the smaller 1950 sample. In particular, [Figure 2](#) shows that the size of the shadow economy in South Asia has been as large as that in Latin America and the Caribbean up to 2000. After that year, South Asia’s shadow-to-formal output ratio nearly doubles: it measures about 70% of formal output in 2015. For comparison, the value for Latin America and the Caribbean is 50%

---

<sup>9</sup> Details on how we classify countries within each dimension are found in [Appendix A](#).

<sup>10</sup> To ease comparison between samples, [Figures 1 and 2](#) share the same  $y$ -axis scale.

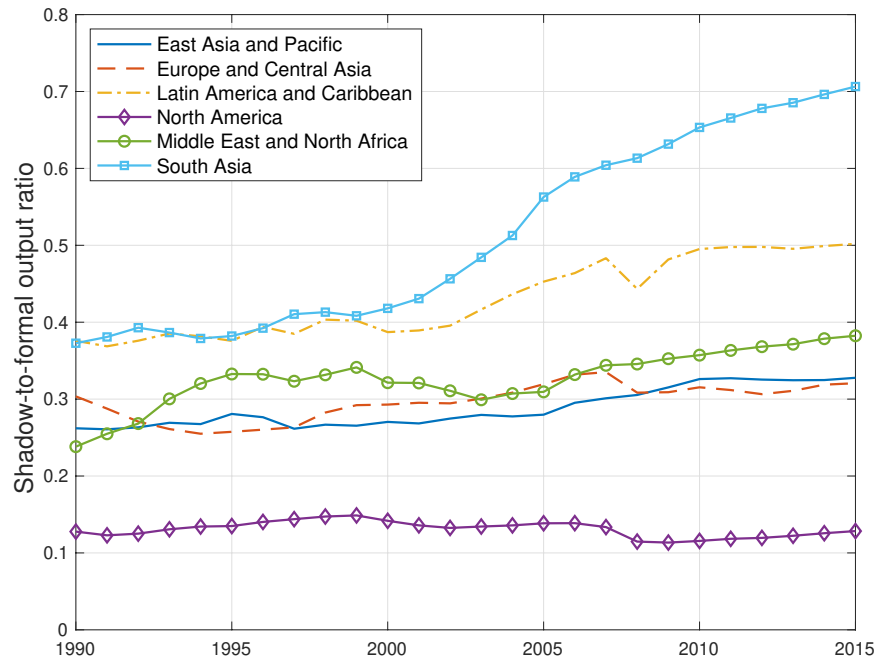


**Figure 1** Average shadow-to-formal ratio, grouped by continent, for countries with data available from 1950.

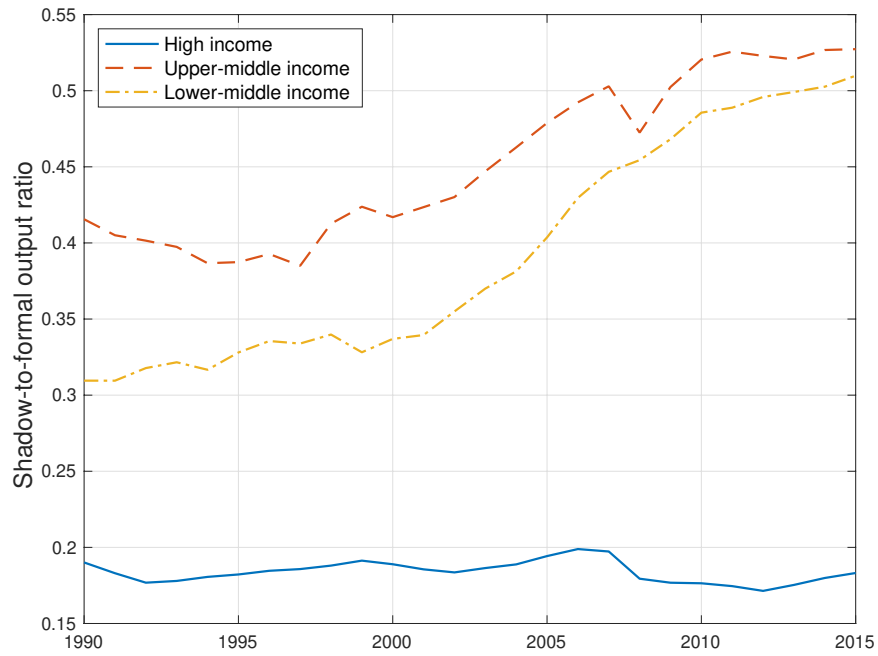
for the same year. For East Asia and Pacific, Europe and Central Asia, and Middle East and North Africa, we observe a slow increase through 1990–2015; the first two regions average a bit over 30% of formal output by 2015, while the last region is closer to 40% for the same year.

## 4.2 By income level

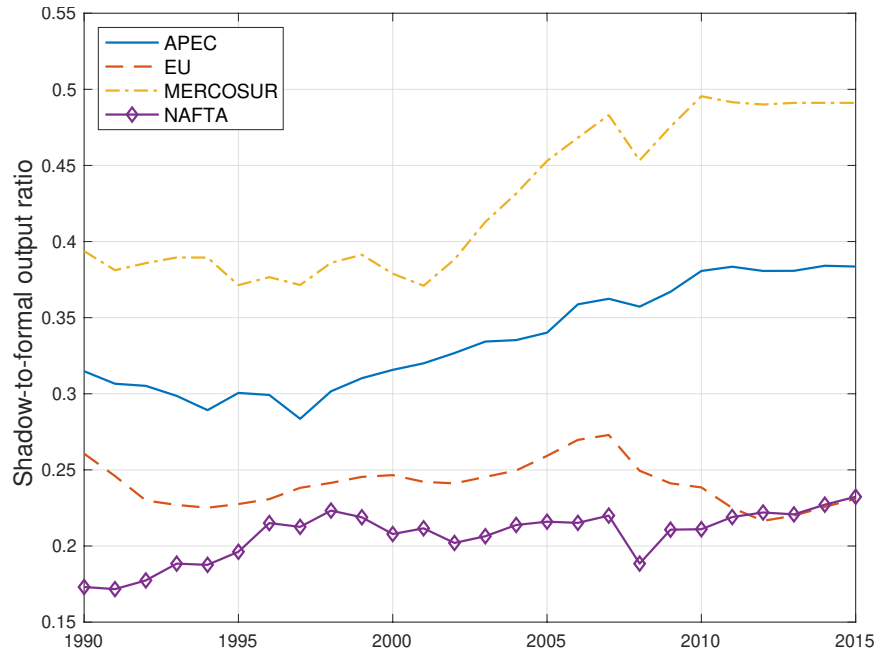
The series in Figure 3 reveal marked disparities among income levels. While the size of the shadow economy is virtually constant over the observed period in high income countries, it is consistently increasing in lower-middle and upper-middle income countries. Note, interestingly, that the shadow-to-formal output ratio is larger in upper-middle income countries—though the gap seems to be closing over time.



**Figure 2** Average shadow-to-formal ratio, grouped by continent, for countries with data available from 1990.



**Figure 3** Average shadow-to-formal ratio, by income per capita, for countries with data available from 1990.



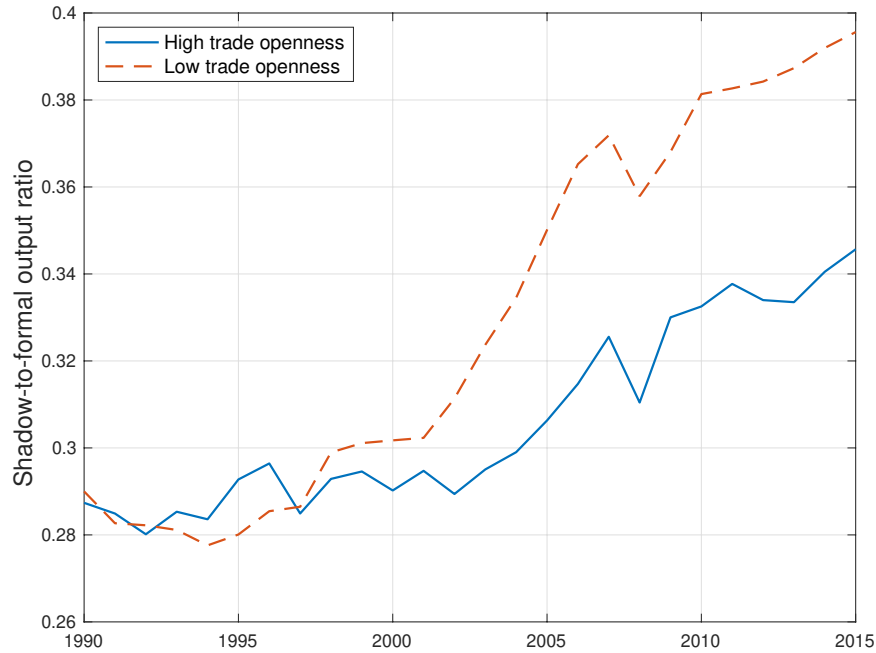
**Figure 4** Average shadow-to-formal ratio, by trade block, for countries with data available from 1990.

### 4.3 By trade block

Figure 4 in large part confirms the trends observed in Figure 2. The size of the shadow economy among MERCOSUR countries—all located in the Latin America and the Caribbean region—is once again the largest relative to the other trade blocks. Countries in the APEC block find their shadow sector steadily increasing in size (relative to formal output) after 1995, while those in the EU exhibit a slight decrease after 2007. Note also how the NAFTA block compares to the North America region from Figure 2: the jump from a bit over 10% to over 20% comes exclusively by the inclusion of Mexico (included in the Latin America and the Caribbean region).

### 4.4 By trade openness

To derive Figure 5, we first calculate the median trade openness for all the countries in the sample. Countries whose trade openness was greater than the median were classified as high trade openness, and vice versa. Figure 5 shows that prior to 2000 both groups of countries had a very similar level



**Figure 5** Average shadow-to-formal ratio, by trade openness, for countries with data available from 1990.

of shadow-to-formal output ratio; a gap begins to form after that year and by 2015, the size of the gap is about 6 percentage points.

## 5 Conclusion

In this paper, we have used the methodology of [Solis-Garcia and Xie \(2018\)](#) to derive estimates of the size and monetary value of the shadow economy for a panel of countries between 1950 and 2015. We also provide some graphical illustrations that characterize the evolution of the size of the shadow economy (relative to formal output) over time. Though far from a careful econometric analysis, these illustrations are an early contribution into tapping the potential of the dataset.

## References

Ceyhun Elgin and Oguz Öztunalı. Shadow economies around the world: Model based estimates. Working Paper 2012/05, Boğaziçi University, Department of Economics, 2012.



Robert Feenstra, Robert Inklaar, and Marcel Timmer. The next generation of the Penn World Table. *American Economic Review*, 105(10):3150–82, 2015.

Jane Ihrig and Karine Moe. Lurking in the shadows: The informal sector and government policy. *Journal of Development Economics*, 73(2):541–77, 2004.

Leandro Medina and Friedrich Schneider. Shadow economies around the world: New results for 158 countries over 1991–2015. Working Paper 6430, Center for Economic Studies & IFO Institute, 2017.

Friedrich Schneider, Andreas Buehn, and Claudio Montenegro. Shadow economies all over the world: New estimates for 162 countries from 1999 to 2007. Policy Research Working Paper 5356, World Bank, 2010.

Mario Solis-Garcia and Yingtong Xie. Measuring the size of the shadow economy using a dynamic general equilibrium model with trends. *Journal of Macroeconomics*, 56:258–75, 2018.

## A Country categories

### A.1 Figures 1 and 2

We follow the region classification offered by the World Bank.<sup>11</sup>

**East Asia and Pacific** Australia, Hong Kong, Indonesia, Japan, Malaysia, New Zealand, Philippines, Singapore, South Korea, Taiwan, Thailand, and Vietnam.

**Europe and Central Asia** Austria, Belgium, Bulgaria, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Latvia, Lithuania,

---

<sup>11</sup> The classification can be found in <https://datahelpdesk.worldbank.org/knowledgebase/articles/906519>.

Luxembourg, Netherlands, Norway, Poland, Portugal, Romania, Russian Federation, Slovak Republic, Slovenia, Spain, Sweden, Switzerland, Turkey, and the United Kingdom.

**Latin America and the Caribbean** Argentina, Brazil, Chile, Colombia, Costa Rica, Jamaica, Mexico, Peru, Uruguay, and Venezuela.

**Middle East and North Africa** Israel and Malta.

**North America** Canada and the United States.

**South Asia** Bangladesh, Pakistan, and Sri Lanka.

## A.2 Figure 3

We follow the income classification offered by the World Bank. (See [Footnote 11.](#))

**Lower-middle income** Bangladesh, Indonesia, Pakistan, Philippines, and Vietnam.

**Upper-middle income** Argentina, Brazil, Bulgaria, Colombia, Costa Rica, Jamaica, Malaysia, Mexico, Peru, Romania, Russian Federation, Sri Lanka, Thailand, Turkey, and Venezuela.

**High income** Australia, Austria, Belgium, Canada, Chile, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hong Kong, Hungary, Iceland, Ireland, Israel, Italy, Japan, Latvia, Lithuania, Luxembourg, Malta, Netherlands, New Zealand, Norway, Poland, Portugal, Singapore, Slovak Republic, Slovenia, South Korea, Spain, Sweden, Switzerland, Taiwan, United Kingdom, United States, and Uruguay.

## A.3 Figure 4

We assign countries to one (or more) of the major trade blocks, depending on membership. We consider the Asia-Pacific Economic Cooperation (APEC), European Union (EU), MERCOSUR, and North American Free Trade Agreement (NAFTA) groups.

**APEC** Australia, Canada, Chile, Hong Kong, Indonesia, Japan, Malaysia, Mexico, New Zealand, Peru, Philippines, Russian Federation, Singapore, South Korea, Taiwan, Thailand, and the United States.

**EU** Austria, Belgium, Bulgaria, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Poland, Portugal, Romania, Slovak Republic, Slovenia, Spain, and Sweden.

**MERCOSUR** Argentina, Brazil, Chile, Colombia, Peru, Uruguay, and Venezuela.

**NAFTA** Canada, Mexico, and United States.

#### A.4 Figure 5

We first obtain a measure of trade openness from the World Bank;<sup>12</sup> then, we calculate the median value of openness (equal to 0.7850) and classify countries into high and low trade openness groups accordingly.

**High trade openness** Austria, Belgium, Bulgaria, Cyprus, Czech Republic, Denmark, Estonia, Finland, Germany, Hong Kong, Hungary, Iceland, Ireland, Jamaica, Latvia, Lithuania, Luxembourg, Malaysia, Malta, Mexico, Netherlands, Poland, Portugal, Romania, Singapore, Slovak Republic, Slovenia, South Korea, Sweden, Switzerland, Thailand, and Vietnam.

**Low trade openness** Argentina, Australia, Bangladesh, Brazil, Canada, Chile, Colombia, Costa Rica, France, Greece, Indonesia, Israel, Italy, Japan, New Zealand, Norway, Pakistan, Peru, Philippines, Russian Federation, Spain, Sri Lanka, Taiwan, Turkey, United Kingdom, United States, Uruguay, and Venezuela.

---

<sup>12</sup> See <http://data.worldbank.org>. We use the Trade (% of GDP) measure, defined as the sum of exports and imports as a percentage of GDP. This measure is attached to a particular year, so we only have one observation.

## B Size of the shadow economy

Country	Start year	Average	Standard deviation	Minimum	Maximum
Argentina	1950	0.2966	0.0474	0.1689	0.3672
Australia	1950	0.1819	0.0213	0.1451	0.2166
Austria	1950	0.1144	0.0245	0.0902	0.1644
Bangladesh	1970	0.4871	0.1975	0.3122	0.9203
Belgium	1950	0.3192	0.0807	0.2175	0.4970
Brazil	1950	0.3789	0.0502	0.3095	0.4535
Bulgaria	1990	0.4483	0.0696	0.3480	0.6114
Canada	1950	0.2233	0.0365	0.1694	0.2882
Chile	1970	0.1949	0.0585	0.1036	0.2827
Colombia	1950	0.3654	0.1140	0.2421	0.6335
Costa Rica	1990	0.4240	0.0890	0.2864	0.5716
Cyprus	1950	0.2544	0.0863	0.1477	0.3981
Czech Republic	1990	0.1951	0.0126	0.1768	0.2163
Denmark	1950	0.2672	0.0686	0.1859	0.4198
Estonia	1990	0.2589	0.0384	0.1962	0.3317
Finland	1950	0.3320	0.0969	0.1859	0.4802
France	1950	0.2414	0.1167	0.1313	0.5292
Germany	1950	0.2510	0.1160	0.1278	0.4764
Greece	1970	0.3335	0.0455	0.2505	0.4486
Hong Kong	1970	0.1828	0.0167	0.1503	0.2170
Hungary	1990	0.3264	0.0288	0.2777	0.4089
Iceland	1950	0.1704	0.0224	0.1322	0.2442
Indonesia	1970	0.2228	0.0625	0.1147	0.3258
Ireland	1950	0.3192	0.0908	0.2012	0.5444

Country	Start year	Average	Standard deviation	Minimum	Maximum
Israel	1990	0.3272	0.0374	0.2318	0.3814
Italy	1950	0.3546	0.0836	0.2629	0.5438
Jamaica	1990	0.3564	0.0327	0.3129	0.4361
Japan	1950	0.1009	0.0226	0.0647	0.1409
Latvia	1990	0.1879	0.0326	0.1447	0.2815
Lithuania	1990	0.1903	0.0303	0.1464	0.2500
Luxembourg	1950	0.1271	0.0412	0.0767	0.2201
Malaysia	1970	0.3861	0.1003	0.2194	0.5930
Malta	1970	0.3099	0.0320	0.2446	0.4005
Mexico	1950	0.2986	0.1305	0.1393	0.5362
Netherlands	1950	0.1978	0.0421	0.1315	0.2840
New Zealand	1950	0.2189	0.0248	0.1607	0.2680
Norway	1950	0.2907	0.0706	0.2213	0.5129
Pakistan	1970	0.4259	0.0431	0.3596	0.5027
Peru	1950	0.6614	0.1939	0.4324	1.0248
Philippines	1970	0.5150	0.0321	0.4154	0.5941
Poland	1990	0.3314	0.0304	0.2705	0.3782
Portugal	1950	0.2200	0.0425	0.1589	0.3142
Romania	1990	0.2288	0.0687	0.1529	0.3957
Russian Federation	1990	0.5013	0.0994	0.3358	0.6262
Singapore	1970	0.1294	0.0494	0.0444	0.2097
Slovak Republic	1990	0.1717	0.0169	0.1370	0.2169
Slovenia	1990	0.2539	0.0250	0.2232	0.3444
South Korea	1970	0.2605	0.0390	0.1621	0.3305
Spain	1950	0.4069	0.1198	0.1989	0.6302
Sri Lanka	1970	0.6387	0.1488	0.3324	0.8740

Country	Start year	Average	Standard deviation	Minimum	Maximum
Sweden	1950	0.2573	0.0492	0.2012	0.3812
Switzerland	1950	0.0926	0.0109	0.0773	0.1140
Taiwan	1970	0.2791	0.0393	0.1801	0.3373
Thailand	1970	0.4755	0.1040	0.2629	0.6439
Turkey	1950	0.5734	0.3446	0.3166	1.7362
United Kingdom	1950	0.1896	0.0334	0.1439	0.2620
United States	1950	0.1043	0.0163	0.0772	0.1371
Uruguay	1990	0.4622	0.0716	0.3254	0.5637
Venezuela	1950	0.3001	0.0735	0.1963	0.4606
Vietnam	1970	0.1852	0.0468	0.1217	0.2617