

The determinants of informality in Mexico's states*

Sean Dougherty

Octavio Escobar

OECD Economics Department ESG Management School

August 2013

Abstract

Informality has important implications for productivity, economic growth, and inequality of income. In recent years, the extent of informal employment has increased in many of Mexico's states. The differences across states in terms of informal employment can be helpful in explaining disparities in economic growth outcomes. This paper studies the determinants of informal employment using diverging outcomes across states and time to identify causal factors. The results suggest that differences in economic development, the prevalence of microenterprises, the quality of labor skills, the cost to start a business, restrictions on foreign investment, and corruption levels explain differences in informal employment among states.

Keywords: Informal employment, microenterprises, regulatory policy, sub-national policy analysis, Latin America, Mexico.

Code JEL: J21, O17, O54.

*This paper began as a technical background paper for the 2013 OECD Economic Survey of Mexico. It has benefitted from feedback at the ATM and DIAL conferences on development and informality as well as seminars at the OECD and University of Groningen. Helpful comments from Sarra Ben Yahmed (SciencesPo), Alexandre Kolev (ILO/OECD), Gérard Duchêne (UPEC), and Patrick Lenain (OECD) were most appreciated. We would like to thank Miguel Messmacher Linartas for motivating this study.

1 INTRODUCTION

Informality has important implications for productivity, economic growth and inequality of income. In recent years, the extent of informal employment has increased in many of Mexico's states, though highly heterogeneously. The substantial differences across states in terms of informal employment can be helpful in explaining differences in economic growth outcomes. Widespread informality can harm the well-being of the population, potentially through negative effects on economic growth, as well as through its interactions with poverty and inequality (OECD, 2009).

The extent of informal employment varies not only among countries, but also across regions within countries. In Mexico, informal employment represents more than half of total employment – 60 percent using our preferred measure. However, informal employment ranges between 45 percent and 80 percent of total employment across states. This paper seeks to answer the question of what might cause the extent of informal employment to vary across Mexico's states to such a high degree. To our knowledge, this is the first paper that analyses the heterogeneity and determinants of informal employment across Mexico's states.¹

In Latin America, after controlling for country characteristics, informality has been found to affect economic growth negatively (Loayza, Serven, & Sugawara, 2009). While high-productivity employment opportunities have been an important part of overall growth in many Asian countries, in Latin American countries, labour has moved from more productive to less productive activities, including, most notably, towards informality (McMillan & Rodrik, 2012). Informal employment has increased over the past two decades in much of the developing world, including countries with high rates of economic growth (Freeman, 2009). However, theoretical models of developing country labor markets treat informal employment as a second-best outcome (Loayza et al., 2009): individuals accept an informal job when a formal job is not available. Some workers may, however, prefer an informal job to avoid taxes and regulations, and they may make a rel-

¹There have been other studies of informality in Mexico. Most recently, Khamis (2012) examines individual-level data and finds that age, education, marital status and scores in the Raven's test, an ability measure, are significant determinants for the various forms of informality.

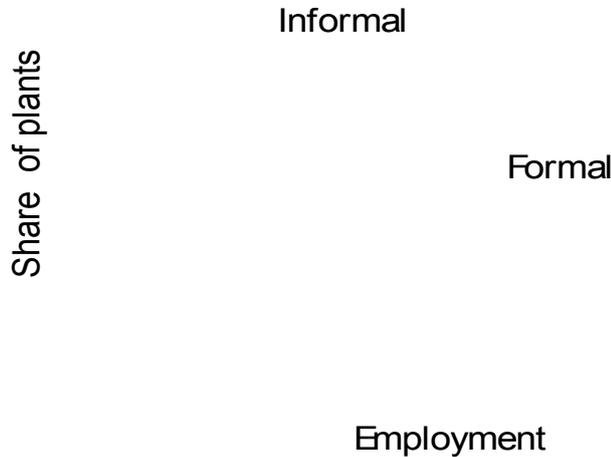
actively better living than with a formal job (Maloney, 2004). Hence, individuals accept an informal job if the benefits of informality outweigh the costs of informality.

Informal employment is strongly linked to firms' productivity. Trying to escape the control of authorities, informal firms remain small, adopt fewer productive technologies, use irregular procurement and divert resources to hide their activities (Dabla-Norris & Inchauste, 2008). For example, informal firms may prefer informal financing since bank financing makes it difficult to mask their activities from authorities. This misallocation of resources harms informal firms' productivity. Since informal firms are generally less productive, aggregate productivity is reduced (Loayza et al., 2009). In the case of Mexico, productivity in the legal informal sector is approximately 50 percent less than in the formal sector (Busso, Fazio, & Levy, 2012). In addition, informal activities use public goods and services without contributing the tax revenue to finance these goods and services. Since public goods and services complement private capital in the process of production, a larger informal sector implies lower productivity growth.

Using panel data from the National Occupation and Employment Survey, we measure informal employment at the sector level for all of Mexico's states over the period 2005 to 2010. Regression analysis is carried out to identify the determinants of this informal employment. The results suggest that differences in economic development, the prevalence of microenterprises, the quality of labor skills, the cost to start a business, restrictions on foreign direct investment and corruption levels explain differences in informal employment among states.

The structure of this paper is as follows: Section 2 presents conceptual and measurement issues related to informal employment. The dynamics of informal employment in Mexico in recent years are discussed in Section 3. Section 4 presents an empirical analysis on the determinants of informal employment, using robust methods that partly address causality. Finally, the conclusions are presented in the last section.

Figure 1: Distribution of employment by firm size in Mexico (2009)



Source: Hsieh and Klenow (2012).

2 DEFINITION OF INFORMAL EMPLOYMENT AND DATA SOURCES

The International Labour Organisation (ILO) has employed two different definitions to study informality. The first one, called the *productivity definition*, measures informality as the informal sector enterprises. These enterprises include family-owned units and microenterprises for which capital accumulation and productivity are very low. The second and more recent definition, called *legalistic definition* focuses on informal employment. The legalistic definition of informality recognizes that informal employment can be found both within and outside of microenterprises. Figure 1 shows that this is the case of Mexico even firms with more than 1,000 employees still employ some informal workers.

Informal employment is defined by the ILO as the number of workers outside the legal framework (*e.g.* jobs for which labor regulations are not applied). In this paper we focus on informal employment rather than on employment in informal firms, which is estimated to be about one-third of employment. We use data on the share of informal employment

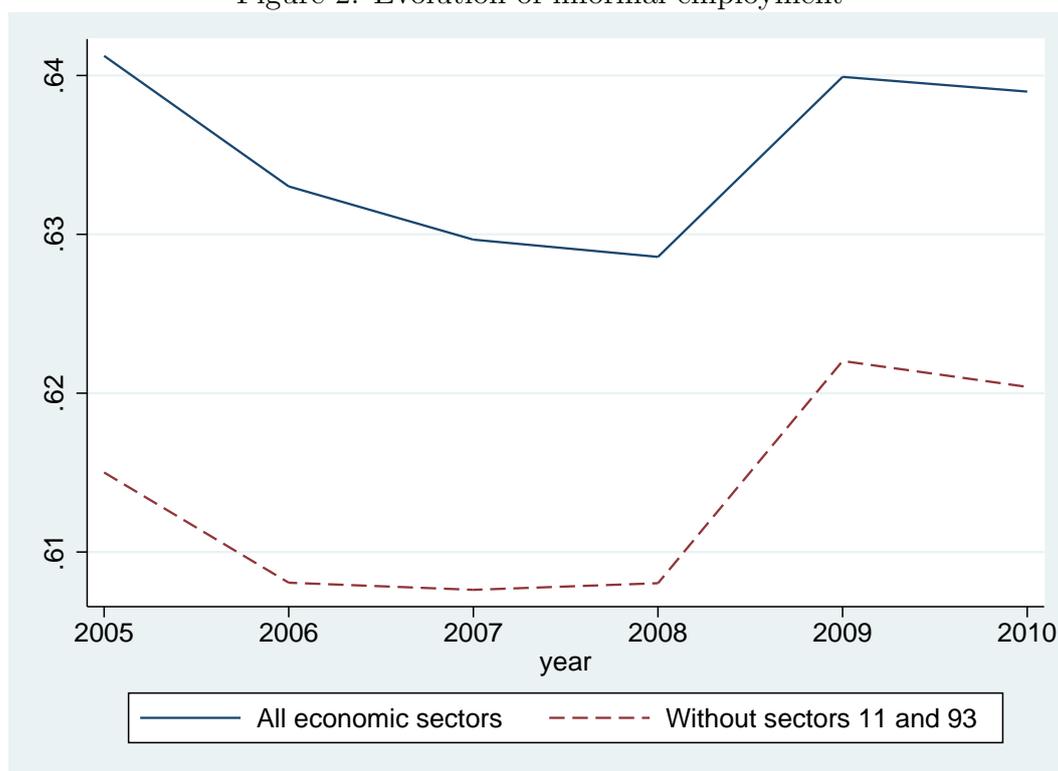
from the National Occupation and Employment Survey (ENOE). The ENOE has polled workers from 120,260 households on a quarterly basis since 2005. This information is then adjusted by INEGI using demographic projections to obtain state and sector-level annual data.

Informal employment is defined as the total number of informal workers, whether employed in formal sector enterprises, informal enterprises or households (ILO, 2003). Informal workers comprise employees and the self-employed which are not subject to national labor legislation, income taxation or social security benefits. As such, informal employment is strongly linked to governmental regulations, not just establishment-level characteristics (*i.e.* scale, legal status, or productivity).

To measure informal employment, we follow the literature for the Mexican case (Binelli & Attanasio, 2010; Bosch & Maloney, 2006; Brandt, 2011), and consider a worker to be informal when he or she does not have access to a health care institution (public or private) granted by his or her workplace. By law, employees must be registered with the national social security agency (IMSS), so if they are not registered, they are then informal according to the ILO (2003) definition. There are some exceptions in the law for state workers who have access to the social security for state workers (ISSSTE) instead of IMSS; other workers can have access to private institutions, military and PEMEX clinics. Note that workers that only have access to *Seguro Popular de Salud* (SPS) are considered as informal workers since, by law, employers cannot propose to access SPS instead of IMSS, ISSSTE, or other institutions. Our measure of informality includes workers without access to one of the health care institutions, excluding the SPS.

Owners and self-employed workers are not required by law to pay social security contributions for themselves. Hence, we are also considering as informal those owners and self-employed without access to health care institutions. These could bias our results; however, most of these persons are self-employed or owners of microenterprises that could be considered to be informal enterprises because of their size (Bosch & Maloney, 2006). Indeed, 97.3 percent of non-primary sector owners and self-employed work for a microenterprise. This can be seen visually in Figure 1, which shows that most employ-

Figure 2: Evolution of informal employment



Source: Calculated using data from the ENOE.

ment in microenterprises (under 10 employees) is informal. Moreover, in our robustness checks, we also use a secondary measure of informality that excludes these workers.

The primary measure of informality that we use, informal employment, is calculated for each state-sector pair in a year as the average of quarterly informal employment. Two economic sectors are excluded from the sample: Agriculture, Forestry, Fishing and Hunting (NAICS code 11) and Public Administration (NAICS code 93). The former has a particular fiscal regime which encourages informality. Indeed, there is no obligation to declare personal income for most small and medium-sized enterprises in the primary sector. In addition, around 40 percent of employees in this sector are owners or self-employed workers. Public Administration is not included since there is little scope for workers in this sector to be informal.

Figure 2 illustrates the evolution of informal employment between 2005 and 2010 with and without the excluded sectors. Without including the primary and public sectors, in-

formal employment is lower. However, the evolution across time is very similar. Informal employment was gradually decreasing until 2007, but it has increased considerably in 2009 following the global financial crisis. Informal employment fell more before and has risen faster since the crisis once the primary sector and public administration are excluded.

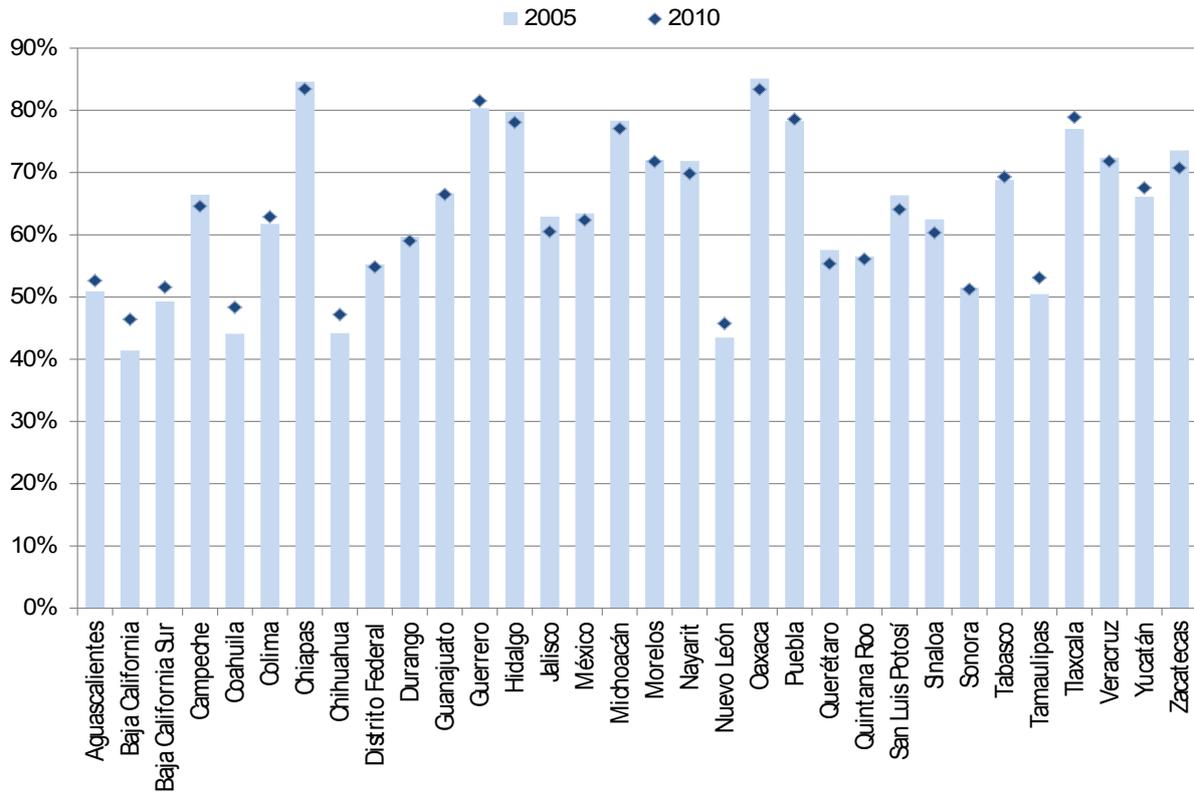
3 DYNAMICS OF EMPLOYMENT INFORMALITY

Figure 3 illustrates the considerable differences across Mexico's states in terms of employment informality. Northern border states (Baja California, Coahuila, Chihuahua, Nuevo León, Sonora, and Tamaulipas), which are the richest states, have the lowest informality rates. Among these states, employment informality fell only in Sonora between 2005 and 2010. The least developed states (Chiapas, Guerrero, Michoacán and Oaxaca) have the highest informality rates, and among these states, employment informality only rose in Guerrero. Even as Mexico's overall informality rose between 2005 and 2010, informality decreased in 11 states. Among these states, Jalisco, Sinaloa, and Zacatecas experienced the greatest reductions in their informality rates between 2005 and 2010.

The period 2005 to 2010 was marked by the financial crisis, and this influenced developments in informality. One-third of the states had GDP growth rates higher than 4 percent between 2005 and 2007. Yet almost two-thirds (20 of 32) of Mexico's states had negative GDP growth rates during the period 2008–2010. Except for Morelos, the growth rates are lower during the crisis period for all states. Five states lost more than seven percentage points in average growth rates. Partly as a consequence, for most states, informality rose faster with the recession, although curiously, in a number of states, employment informality decreased even during the crisis period. These patterns were sustained across most sectors of economic activity, though the overall informality rate increased sharply in manufacturing in 2009, and then fell again sharply in 2010.

There are also important differences in informality rates across economic sectors, as shown by Figure 4. At the high end, the sectors of Construction and Other Services have a rate of informal employment over 80 percent. At the low end, Utilities, Finance and Insurance, and Educational Services have a rate of informal employment under 20 per-

Figure 3: Employment informality by state

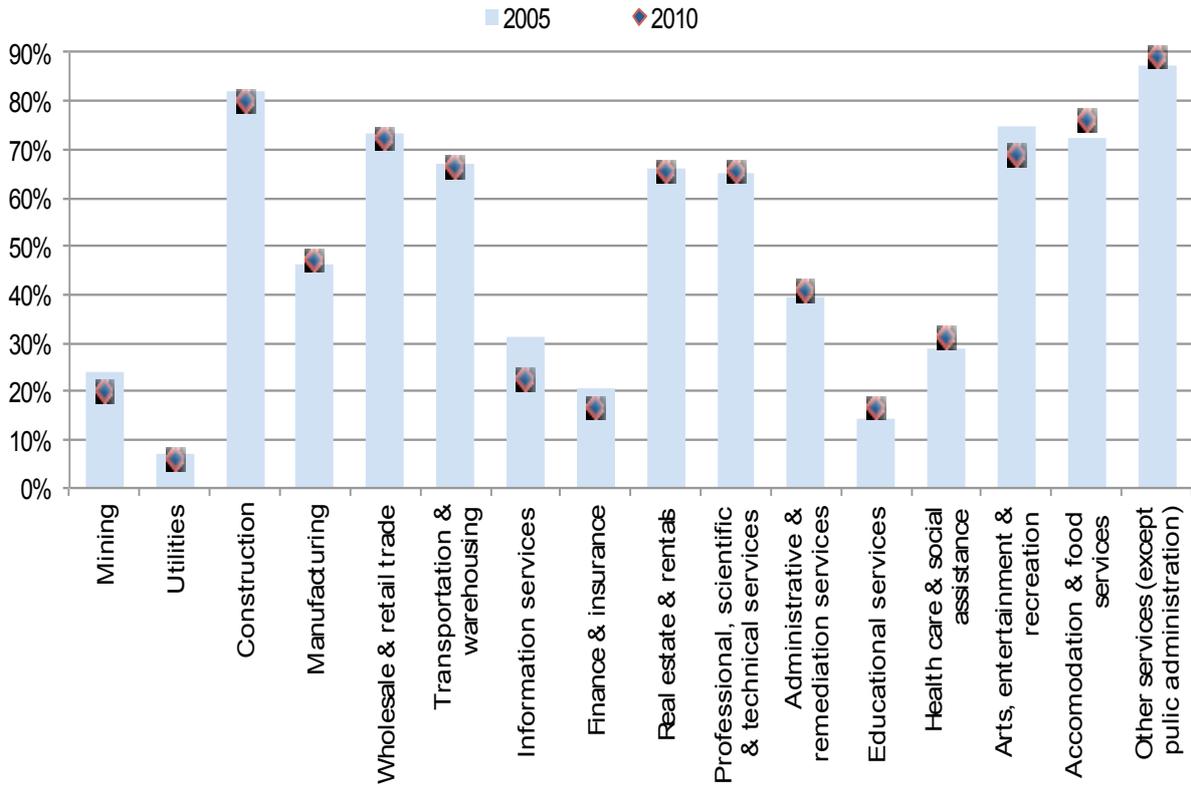


Source: Calculated using data from the ENOE.

cent. Within states, economic recession affected employment informality for virtually all sectors. Employment informality was decreasing before 2008 for most economic sectors, but it increased after 2008. There are, however, differences in the dynamics of informality across economic sectors. Indeed, the Manufacturing and Wholesale and Retail Trade sectors recovered faster from the crisis, with informality starting to decrease in 2010.

Comparing informal employment rates in 2010 to those in 2005 by economic sector, we also observe different patterns. For most of the economic sectors (9 of 16), informal employment decreased. The reduction of informal employment is particularly important for the Information Services sector, which already was one of the sectors with the lowest informality rate. Though the Arts, Entertainment, and Recreation sector is a sectors with one of the highest informality rates, it also had a substantial decrease in informality. We then note that improvements in informality are not exclusive to sectors with low informality rates, neither to those of high informality rates. Figure 4 also shows that informality has increased somewhat more in the services sectors. Informality has increased in sectors

Figure 4: Employment informality by sector



Source: Calculated using data from the ENOE.

with a lot of informality such as Accommodation and Food Services and Other Services, as well as in sectors with low rates of informality such as Educational Services and Health Care and Social Assistance. In the secondary sector, differences in informality between 2010 and 2005 are less important. Among secondary sectors, informal employment has increased only for Manufacturing.

4 DETERMINANTS OF INFORMAL EMPLOYMENT

The previous section illustrated the heterogeneity of employment informality among states and economic sectors. In this section, we conduct an econometric analysis to identify the determinants of informal employment at the state and sector levels.

(a) Variables definition and data sources

For each state-sector pair, employment informality is calculated as the workers without access to health care in percent of total workers. Employment informality is calculated for the 31 Mexican states and the Federal District for the following economic sectors (NAICS code in parentheses): (21) Mining, (22) Utilities, (23) Construction, (31-33) Manufacturing, (43-46) Wholesale and Retail Trade, (48-49) Transportation and Warehousing, (51) Information Services, (52) Finance and Insurance, (53) Real Estate and Rental and Leasing, (54) Professional, Scientific, and Technical Services, (56) Administrative and Support, Waste Management, and Remediation Services, (61) Educational Services, (62) Health Care and Social Assistance, (71) Arts, Entertainment, and Recreation, (72) Accommodation and Food Services, and (81) Other Services (except Public Administration).

Theoretical and empirical literature suggests that taxes and social security burdens, government effectiveness, the stringency of regulations and official economy influence informality (Schneider, Buehn, & Montenegro, 2010). These factors would influence both formal labor demand and formal labor supply decisions.

(i) The economic environment

A favourable economic environment is likely to contribute to a decrease in employment informality. Formal employment opportunities are more widespread in a growing economy than in an economy in recession (Loayza et al., 2009). To control for this effect, we use data on GDP per capita, the inflation rate and the unemployment rate at the state level from the INEGI.

In addition, the presence of multinational enterprises influences the labor market (Markusen & Venables, 1999). Hence, we calculate the stock of FDI as a share of GDP for each state using data from the *Secretaría de Economía* and INEGI. The stock of FDI is calculated using the perpetual inventory method: following Escobar (2012), FDI stocks are defined as annual FDI inflows minus the accumulated consumption of FDI.

Informal firms may prefer informal financing since bank financing makes it difficult

to mask their activities from authorities (Dabla-Norris & Inchauste, 2008). Hence, firms make a choice between the costs of becoming formal and the costs of financing their activities. Interest rates are the same across Mexico’s states, however there are differences in the access to credit by state which influence the cost of financing among states. To control for these differences, we use Banxico’s data on commercial bank credit as a share of GDP by state.

(ii) Sector specialisation, skilled-labor intensity, and the size of the firms

To control for industrial heterogeneity among sectors and across states, we construct a variable of sectoral specialisation following the Balassa (1965) index using output data from INEGI. The specialisation of state i in sector j is calculated as the sector output as a share of state’s output divided by the sector j output as a share of country M total output:

$$S_{i,j} = \frac{output_{i,j}/output_i}{output_{M,j}/output_M} \quad (1)$$

If the coefficient is higher than one then the state i is more specialised in sector j than the average of states in the country. On the contrary, if the coefficient is lower than one, the specialisation of the state in sector j is weakest than the specialisation of the average state in sector j .

The size of the firm influences the likelihood for use of informal labor. ILO (2003) and Bosch and Maloney (2006) point out that small scale is an important characteristic of informal enterprises, particularly of family-based ones. Hence, we measure the intensity of microenterprises as the share of labor in microenterprises (under 10 employees) at the state-sector level. This variable is computed using data from the ENOE.

Skilled labor may be an important factor in and of itself in allowing workers to participate in larger, more capital-intensive enterprises. In addition, skilled-labor intensity increases the power of negotiation since this factor is relatively scarce in the country. Hence, education may increase the likelihood of finding a formal job. For each state-sector pair, skilled-labor is measured as the share of workers with at least high school

using data from the ENOE².

(iii) Tax and social security contribution burdens

Tax and social security contribution burdens are thought by many to be among the main causes of informality (Levy, 2010). From an employment supply perspective, workers have incentives to accept an informal job the bigger the difference between before and after-tax earnings. From an employment-demand perspective, social security contribution burdens influence labor costs. Concerning Mexico's states, each state has fiscal autonomy for some taxes. However, social security contribution burdens are the same among the states. To evaluate if differences in taxation are important, we use an indicator of the efficiency of tax administration.³ The efficiency of tax administration is defined as the ratio of the payroll of tax administration to the taxes collected. This index is available from the *Instituto Mexicano para la Competitividad* database.

(iv) Public institutions and intensity of regulations

The stringency of regulations increases the costs of formal labor and may drive workers to informality (Almeida & Carneiro, 2012; Schneider et al., 2010). To measure the stringency of regulations, we employ the cost to start business as a share of income per capita from the World Bank's Sub-national Doing Business data for Mexico. This cost includes all official fees and fees for legal or professional services if these services are required by law. Possible bribes are not included in this cost.

In addition to the stringency of regulations and tax and social security contribution burdens, the efficiency of application of these regulations influences informal employment. Better compliance with mandated benefits makes it attractive to be a formal employee,

²The share of high school graduates measures the quantity of skilled-labor but not the quality. In robustness tests we include measures of the quality of education such as PISA and ENLACE scores, but they were not statistically significant. PISA and ENLACE scores may reflect the quality of future labor force but not necessarily the quality of the present one. Unfortunately, these tests began in Mexico relatively recently and there is limited time-series variation.

³In robustness checks, following Schneider et al. (2010), we used the share of direct and indirect taxes of overall state's income as an alternative measure of taxation, as well as the state government revenue as a percent of GDP. Estimates for these variables were, however, not significant in all specifications and estimators.

even if wages are lower (Almeida & Carneiro, 2012). We employ two different measures to control for regulations' enforcement:⁴

1. Survey-based corruption prevalence from *Transparencia Mexicana*. This index measures the corruption in public services. It uses a sample of around 15,000 households around the country to which they are asked about 35 public services. The index is calculated as $\text{Corruption} = \frac{\text{Number of bribes given to public services}}{\text{Number of times public services were used}}$.
2. The rate of labor dispute resolution using labor court data from INEGI. Yearly data on individual and collective disputes are used, as well as on the number of individual and collective cases resolved. We compute the rate of labor dispute resolution as the number of cases resolved (both individual and collective) in the year as a share of new cases in that year.

(b) Empirical results

In order to examine the determinants of employment informality set out in the previous section, we estimate the following equation for state i , sector j , and time t :

$$\ln(\text{Informality}_{i,j,t}) = \alpha + \beta \ln(X_{i,j,t}) + u_i + u_j + u_t + \varepsilon_{i,j,t} \quad (2)$$

where X is the matrix of independent variables; u_i , u_j , and u_t are the specific characteristics of state i , sector j and time t , respectively; and $\varepsilon_{i,j,t}$ is the error term.

Table 1 presents the estimates for the determinants of informal employment. Columns 1 and 2 show the results using the Ordinary Least Squares (OLS) and Least Square Dummy Variable (LSDV) estimators. Both regressions use Heteroskedastic and Autocorrelation-Consistent HAC estimates⁵.

⁴We employed alternative variables such as Moody's index of enforceability of commercial contracts, and the law enforcement and security perception index from *Instituto Mexicano para la Competitividad*. Results are in the same direction to those found using corruption prevalence, though the conceptual relationship with informality is less straightforward.

⁵Applying the Wooldridge (2002) test for serial correlation in panel data, the null hypothesis of no serial correlation is rejected. Hence, we report HAC standard errors. We also estimate regressions using standard errors robust to heteroskedasticity and intra-group correlation and we did not find significant differences.

Table 1: Estimates for the determinants of informal employment

Dependent variable: Informality share						
	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	LSDV	GMM	GMM- FE	GMM	GMM- FE
GDP per capita	-0.091** (0.035)	-0.035 (0.151)	-0.069* (0.037)	-0.443 (0.285)	-0.073* (0.039)	-0.444 (0.280)
Inflation rate	0.007 (0.010)	-0.005 (0.009)	0.008 (0.010)	-0.005 (0.008)	0.006 (0.010)	-0.006 (0.008)
Unemployment rate	-0.015 (0.044)	0.073** (0.036)	0.050 (0.056)	0.029 (0.093)	0.057 (0.055)	0.027 (0.090)
Stock of FDI as a share of GDP	-0.015* (0.008)	-0.006 (0.006)	-0.047** (0.023)	0.035** (0.015)	-0.044** (0.022)	0.037** (0.016)
Commercial bank credit as a share of GDP	-0.006 (0.027)	-0.002 (0.025)	0.032 (0.035)	0.180* (0.097)	0.029 (0.034)	0.191** (0.096)
Sector specialisation	0.005 (0.028)	-0.133*** (0.020)	-0.004 (0.031)	-0.124*** (0.027)	-0.005 (0.032)	-0.119*** (0.029)
Share of high school graduates	-0.135*** (0.026)	-0.203*** (0.042)	-0.101*** (0.032)	-0.316*** (0.098)	-0.100*** (0.030)	-0.333*** (0.099)
Efficiency of tax administration	-0.132*** (0.049)	0.043 (0.047)	-0.141** (0.057)	-0.012 (0.058)	-0.149*** (0.057)	-0.018 (0.060)
Corruption prevalence	0.130*** (0.034)	-0.033 (0.020)	0.129*** (0.036)	-0.012 (0.032)	0.133*** (0.035)	-0.012 (0.031)
Rate of labour dispute resolution	0.008 (0.040)	0.003 (0.019)	0.014 (0.034)	0.008 (0.023)	0.021 (0.034)	0.009 (0.023)
Microenterprise share	0.521*** (0.023)	0.284*** (0.030)	0.563*** (0.028)	0.338*** (0.045)	0.275* (0.147)	0.105 (0.087)
Cost to start a business as a share of income	-0.135*** (0.038)	-0.071* (0.039)	-0.137*** (0.043)	-0.157** (0.079)	0.019 (0.063)	-0.032 (0.071)
Microenterprise share × cost to start a business					0.112* (0.057)	0.095*** (0.034)
Observations	3063	3063	2550	2550	2550	2550
R-squared	0.754	0.884	0.741	0.874	0.744	0.874
RMSE	0.430	0.298	0.448	0.312	0.446	0.313
Hansen J <i>p</i> -value			0.519	0.858	0.388	0.844

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. All variables are expressed in log form. The estimators are: OLS - ordinary least squares; LSDV - least squares dummy variables; GMM - generalized method of moments. Heteroskedastic and auto-correlation robust standard errors are in parentheses. Each regression includes a constant and time dummies not reported here. LSDV and GMM-FE also include sector and state dummies not reported here. Hansen J-test reports the p -values for the null hypothesis of instruments validity.

OLS estimates (column 1) suggest that the GDP per capita, the stock of FDI, the share of high school graduates, the efficiency of tax administration, the prevalence of corruption, the microenterprise share, and the cost to start a business are significant determinants of informal employment in Mexico. On the one hand, GDP per capita, the stock of FDI, the share of high school graduates, the efficiency of tax administration, and the cost to start a business reduce employment informality. On the other hand, the

microenterprise share and prevalence of corruption increase employment informality.

Only the microenterprise share, the share of high school graduates, and the cost to start a business are still significant when controlling for state and sector-specific characteristics (column 2). In addition, LSDV estimates show that informality rises with unemployment; while informality decreases with sector specialisation. The R-squared and root means-squared error (RMSE) values suggest that LSDV estimator performs better than the OLS estimator, highlighting the heterogeneity among Mexico's states and economic sectors. In addition, coefficient values for many sectors and states dummies are significant at the 5 percent level.

Endogeneity issues arise in estimating Equation 2. First, there is potential simultaneity among independent variables because of possible reverse causality, which could result in the endogeneity of some regressors. For example, unemployment may influence employment informality, but employment informality could also reduce unemployment. Second, there is also potential endogeneity because of omitted variables where data is not available. These omitted variables are captured by the unobserved state and sector-specific characteristics (fixed effects), but may be correlated with some explanatory variables.

We start by identifying the endogenous variables. We apply the Wu-Hausman F and the Durbin-Wu-Hausman Chi-squared tests for each explanatory variable. We reject the null of exogeneity for the following variables: the unemployment rate, stock of FDI, sector specialisation, high school graduates, microenterprise share and cost to start a business. Even if the Wu-Hausman F test p-values are sufficiently large not to reject the null of exogeneity for the GDP per capita variable, we assume this variable as endogenous under theoretical arguments. A solution to endogeneity issues is the two-step feasible generalised method of moments GMM estimator. We estimate the GMM estimator using one-year lagged values of endogenous and exogenous explanatory variables as instruments.

With the default GMM estimator (column 3), results are quite similar to those presented previously. However, informality is less sensitive to GDP per capita and more sensitive to FDI. When controlling for fixed effects (FE) (column 4), there are more differences compared to LSDV estimates. Thus, FDI and commercial bank credit be-

come significant, while unemployment becomes insignificant. This is likely due to multicollinearity, as sectoral specialisation, high school graduates, and cost to start a business are all negatively correlated with unemployment. Finally, the Hansen J -statistic test for over-identifying restrictions is employed to evaluate the validity of a group of instruments when estimators are adjusted to heteroskedasticity and autocorrelation (Roodman, 2009). The p -value suggest that the instruments are valid.

Only the share of high school graduates, microenterprise share, and cost to start a business are significant using all of the OLS, LSDV, GMM, and GMM-FE estimators. To reduce labor informality it is necessary to promote education, as well as to reduce the share of microenterprises. However, results also suggest that increasing the cost to start a business leads to a reduction of informality. The following two arguments can be helpful to explain this result. First, increasing the entry costs limits the entrance of less productive firms (Melitz, 2003) which employ informal workers. Second, increasing the cost of starting a business also limits entrepreneurship (Stel, Storey, & Thurik, 2007); since some people become entrepreneurs simply to employ themselves and also because owners are not obligated to pay social security contributions for themselves, restricting entrepreneurship can limit informal employment. Hence, reducing the ease of doing business for the smallest firms may have a negative impact through the proliferation of microenterprises.

A better regulatory framework increases the size of existing enterprises (Dabla-Norris & Inchauste, 2008; Dougherty, 2013). As a measure of the stringency of regulations, increasing the cost to start a business may influence the growth of microenterprises and indirectly informality. Hence, the effects of costs to start a business may depend on the microenterprise share. We thus add to the model an interaction between the variables microenterprise share and cost to start a business. By interacting microenterprise share with cost to start a business, we let the marginal impact of cost to start a business be a linear function of microenterprise share, and *vice versa*.

Columns 5 and 6 present the results for the specification that includes the interaction between the microenterprise share and the cost to start a business variables. Results

show that the cost to start a business variable alone becomes insignificant. The estimates with the interaction variable show that the more the microenterprise share is important, increasing the cost to start a business increases of informality. We can also interpret these results as the impact of increasing microenterprises on informality is more important when the cost to start a business is higher.

We consider owners and self-employed without access to health care institutions as informal workers, and this could bias our results since these workers are not required by law to pay social security contributions for themselves. To avoid any possible bias of considering some owners and self-employed workers as informal labor, we exclude these workers from the dependent variable. More precisely, we replace the informal employment share by the informal employees share as the dependent variable. The informal employees share is measured as the share of employees without access to a health care institution granted by their workplace.

Table 2 shows the GMM estimates for the determinants of informal employees rate. Without controlling for fixed effects (column 1), the share of informal employees decreases with GDP per capita, FDI, high school graduates, efficiency of tax administration, and cost to start a business, and it increases with unemployment, corruption, and microenterprises share. These results are similar to those of informal employment. The informal employees rate is however more sensitive to unemployment, GDP per capita, FDI, and high school graduates.

For some variables, the significance of estimates differs when controlling for states and sectors' heterogeneity (column 2). The stock of FDI is still significant, but its sign changes. GDP per capita, unemployment, efficiency of tax administration, corruption, and cost to start a business lose their significance. On the other hand, commercial bank and sector specialisation become significant. These results are also similar to those of informal employment, but the informal employees rate is less sensitive to changes in the cost to start a business which is not significant when controlling for state and sector heterogeneity. However, when introducing the interaction between the variables microenterprise share and cost to start a business (columns 3 and 4), the marginal effects

of microenterprise share, cost to start a business, and the interaction term on informal employees are quite similar to those obtained using informal employment.

Table 2: Estimates for the determinants of informal employees

Dependent variable: Informal employees share				
	(1)	(2)	(3)	(4)
	GMM	GMM- FE	GMM	GMM- FE
GDP per capita	-0.138*** (0.043)	-0.453 (0.339)	-0.142*** (0.044)	-0.462 (0.336)
Inflation rate	0.005 (0.011)	-0.012 (0.009)	0.002 (0.011)	-0.012 (0.009)
Unemployment rate	0.121* (0.064)	0.147 (0.105)	0.129** (0.064)	0.143 (0.103)
Stock of FDI as a share of GDP	-0.085*** (0.024)	0.040** (0.017)	-0.082*** (0.023)	0.042** (0.017)
Commercial bank credit as a share of GDP	0.035 (0.039)	0.207** (0.104)	0.033 (0.039)	0.215** (0.104)
Sector specialisation	-0.011 (0.035)	-0.139*** (0.031)	-0.011 (0.036)	-0.135*** (0.032)
Share of high school graduates	-0.207*** (0.036)	-0.325*** (0.111)	-0.205*** (0.034)	-0.342*** (0.111)
Efficiency of tax administration	-0.165*** (0.063)	0.007 (0.060)	-0.171*** (0.063)	-0.002 (0.062)
Corruption prevalence	0.186*** (0.043)	0.008 (0.036)	0.190*** (0.042)	0.008 (0.036)
Rate of labour dispute resolution	0.005 (0.039)	0.017 (0.028)	0.011 (0.039)	0.017 (0.028)
Microenterprise share	0.498*** (0.027)	0.346*** (0.047)	0.194 (0.143)	0.148* (0.088)
Cost to start a business as a share of income	-0.211*** (0.052)	-0.115 (0.087)	-0.046 (0.072)	-0.010 (0.081)
Microenterprise share \times Cost to start a business			0.119** (0.057)	0.080** (0.035)
Observations	2545	2545	2545	2545
R-squared	0.672	0.848	0.674	0.848
RMSE	0.510	0.347	0.508	0.347
Hansen J p -value	0.244	0.675	0.181	0.701

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. All variables are expressed in log form. The estimators is the GMM - generalized method of moments. GMM-FE also include sector and state dummies not reported here. Heteroskedastic and auto-correlation robust standard errors are in parentheses. Each regression includes a constant and time dummies not reported here. GMM-FE also includes sector and state dummies not reported here. Hansen J-test reports the p -values for the null hypothesis of instruments validity.

Improving the ease of doing business may have negative or positive effects on informality. First, a higher cost to start a business limits entrepreneurship and thus the creation of self-employment, which is usually informal. It also limits the creation of less productive firms which may employ informal workers. Second, reducing the cost to start a business

reduces the cost of formal employment. In our baseline results, if the reduction of the cost to start a business leads to creation of small/medium-sized or large enterprises rather than creation of microenterprises, the positive effect dominates and informality decreases.

According to the productivity definition of informality, microenterprises are strongly related to the informal sector. We can then suspect the endogeneity bias created by this variable are not completely solved by the standard GMM estimator. In addition, estimates presented on Tables 1 and 2 show differences in significance when controlling for and not controlling for fixed effects. Most importantly, sometimes even the sign of the coefficient value shifts such as the stock of FDI and the corruption variable. Bond, Hoeffler, and Temple (2001) show that the OLS estimator is upward-biased and inconsistent while the LSDV estimator is downward-biased and inconsistent. Hence, the real effect must be between estimates without controlling for fixed-effects and estimates controlling for fixed-effects. A solution to these issues is the generalised method of moments estimator (GMM) and the Blundell and Bond (1998) system GMM estimator. The system GMM estimator is derived from the estimation of a system of two simultaneous equations. The first or original equation is expressed in levels, and uses lagged first differences of explanatory variables as instruments. The second, or transformed equation, expressed in first differences, uses lagged levels of explanatory variables as instruments. The system GMM is designed for the situations with few time periods and many individuals, which is the case of our panel. Indeed, our panel dataset has a relatively short time period ($T = 6$ years) and a large cross-sectional state/sector dimension ($N = 525$). This type of panel increases the likelihood of endogeneity because of omitted variables. Normally, shocks to fixed effects decline over time.

To obtain valid system GMM estimators, we need exogenous instruments. With the GMM estimator, the Hansen J -statistic test for over-identifying restrictions is employed to evaluate the validity of a group of instruments. We also evaluate the validity of a subset of instruments using the “difference-in-Hansen” statistics. An estimate of the validity of the instruments is generated for each explanatory variable. The instruments not satisfying the difference-in-Hansen test are excluded. Only for the unemployment variable were the

GMM instruments were not exogenous, in both levels and differences. Hence, we exclude the system GMM instruments of this variable from all system GMM estimates. Finally, we use Windmeijer (2005) finite sample corrected standard errors. This makes two-step estimation is more efficient than the the one-step estimator (Roodman, 2009).

We proceed then to estimate the system GMM estimator (column 1 of Table 3). Time-lagged informal employees share is significant, which supports the dynamic process of informality. GDP per capita, stock of FDI, and share of high school graduates reduce informality; corruption prevalence and microenterprise share increase it. Concerning instrument validity, the Hansen J -test suggests that the instruments used are exogenous. This test can be weakened by instrument proliferation, but note that in our case, fortunately, the number of instruments is far lower than the number of groups. The values reported for AR(1) and AR(2) are the p -values for first and second order autocorrelated disturbances in the first-differenced equation. As expected, there is high first order autocorrelation, and no evidence for significant second order autocorrelation. These test statistics suggest that the specification is appropriate.

We introduce the interaction between microenterprise share and cost to start a business in column 2. Compared to previous results, microenterprises share, cost to start a business, and the interaction term are not significant, though the signs remain the same as in the earlier estimates. The major determinants of informal employees share are then GDP per capita, stock of FDI, share of high school graduates, and corruption prevalence.

We suspect that the endogeneity of GDP per capita variable generates bias, even though the system GMM estimator should handle endogeneity issues. We present results without including the GDP per capita variable (columns 3 and 4). Estimates are similar to those found using GDP per capita variable. The only difference concerns the efficiency of tax administration variable which becomes significant in column 3. This variable apparently captures some of the variance of GDP per capita.

Results reported on Table 3 confirm that to reduce labor informality it is necessary to promote economic development, education, FDI openness, as well as to reduce corruption and facilitate the growth of enterprises.

Table 3: Estimates for the determinants of informal employees using dynamic panel

	Dependent variable: Informal employees share			
	(1)	(2)	(3)	(4)
	System GMM	System GMM	System GMM	System GMM
Year lagged informal employees rate	0.617*** (0.099)	0.619*** (0.100)	0.676*** (0.091)	0.682*** (0.092)
GDP per capita	-0.116** (0.052)	-0.115** (0.053)		
Inflation rate	-0.002 (0.008)	-0.002 (0.008)	-0.000 (0.008)	-0.000 (0.008)
Unemployment rate	0.065 (0.044)	0.066 (0.044)	0.044 (0.047)	0.046 (0.049)
Stock of FDI as a share of GDP	-0.021*** (0.007)	-0.021*** (0.007)	-0.021** (0.008)	-0.021** (0.009)
Commercial bank credit as a share of GDP	-0.005 (0.027)	-0.003 (0.027)	0.018 (0.029)	0.020 (0.029)
Sector specialisation	-0.038 (0.060)	-0.041 (0.059)	-0.007 (0.053)	-0.012 (0.053)
Share of high school graduates	-0.158* (0.092)	-0.158* (0.091)	-0.176* (0.100)	-0.179* (0.103)
Efficiency of tax administration	-0.018 (0.033)	-0.018 (0.033)	-0.063* (0.038)	-0.061 (0.038)
Corruption prevalence	0.043** (0.022)	0.042* (0.022)	0.037* (0.022)	0.035 (0.022)
Rate of labour dispute resolution	-0.021 (0.020)	-0.021 (0.020)	-0.030 (0.021)	-0.032 (0.022)
Microenterprise share	0.157*** (0.045)	0.092 (0.157)	0.134*** (0.047)	0.082 (0.167)
Cost to start a business as a share of income	-0.036 (0.044)	-0.005 (0.070)	-0.020 (0.052)	0.010 (0.069)
Microenterprise share \times cost to start a business		0.026 (0.061)		0.020 (0.065)
Observations	2540	2540	2540	2540
Instruments	114	114	100	100
Groups	512	512	512	512
Hansen J p-value	0.378	0.366	0.151	0.150
AR(1) p -value	0.000	0.000	0.000	0.000
AR(2) p -value	0.860	0.859	0.926	0.931

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. All variables are expressed in log form. The estimator system GMM is the Two-step Blundell and Bond (1998) system GMM. Cluster adjusted (state-sector pair level) robust standard errors are in parentheses. The Windmeijer (2005) finite sample correction for standard errors is employed. Each regression includes a constant and time dummies not reported here. Hansen J-test reports the p-values for the null hypothesis of instruments validity. The p -values reported for AR(1) and AR(2) are the p-values for first and second order autocorrelated disturbances.

5 CONCLUSIONS

Informal employment is known to limit productivity and, as a consequence, economic growth. Understanding the differences in terms of informal employment across Mexico's

states could be helpful to understand the differences in terms of productivity and development. In this paper, we study the determinants of informal employment using state-sector level panel data for the period 2005 to 2010. After controlling for endogeneity and heterogeneity, and for different specifications, a common result emerges: GDP per capita, FDI stocks, skilled-labor intensity, the microenterprise share, and corruption prevalence significantly influence informal employment in Mexico. These results also suggest that there are important differences in terms of these variables among Mexico's states which explain differences in terms of informal employment.

In order to reduce informal employment in Mexico, authorities need to promote economic growth, education and FDI openness. Employment opportunities are more plentiful in a growing economy; hence the likelihood for a worker to find a formal job is also higher. This likelihood is also higher if the worker has completed high school since skilled-labor intensity of the sector reduces the informality rate. The presence of multinational firms contributes to a decrease in informal employment. Multinational firms are larger and tend to follow labor legislation more systematically than domestic ones.

The prevalence of corruption and weak legal institutions leads to an increase in employment informality. Widespread corruption reduces the benefits of accessing to public goods and services, discouraging workers to demand a formal job; it also weakens the strength of controls by the authorities which encourage firms to propose informal jobs to workers. In addition, since better legal system quality increases enterprises' size (Dougherty, 2013), reducing corruption may also help limit the share of microenterprises.

REFERENCES

- Almeida, R., & Carneiro, P. (2012). Enforcement of Labor Regulation and Informality. *American Economic Journal: Applied Economics*, 4(3), 64-89.
- Balassa, B. (1965). Trade Liberalisation and 'Revealed' Comparative Advantage. *The Manchester School*, 33(2), 99-123.
- Binelli, C., & Attanasio, O. (2010). Mexico in the 1990s: the Main Cross-Sectional Facts. *Review of Economic Dynamics*, 13(1), 238-264.

- Blundell, R., & Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, 87(1), 115-143.
- Bond, S. R., Hoeffler, A., & Temple, J. (2001). *GMM Estimation of Empirical Growth Models* (CEPR Discussion Papers No. 3048). C.E.P.R. Discussion Papers.
- Bosch, M., & Maloney, W. (2006). *Gross worker flows in the presence of informal labor markets: the Mexican experience 1987-2002* (Policy Research Working Paper Series). The World Bank.
- Brandt, N. (2011). Informality in Mexico. *OECD Economics Department Working Papers*(896).
- Busso, M., Fazio, M. V., & Levy, S. (2012). *(In)Formal and (Un)Productive: The Productivity Costs of Excessive Informality in Mexico* (IDB Working Paper Series No. 4789). Inter-American Development Bank, Research Department.
- Dabla-Norris, E., & Inchauste, G. (2008). Informality and Regulations: What Drives the Growth of Firms? *IMF Staff Papers*, 55(1), 50-82.
- Dougherty, S. (2013). Legal reform, contract enforcement and firm size in Mexico. *OECD Economics Department Working Papers*(1042).
- Escobar, O. (2012). Foreign direct investment (FDI) determinants and spatial spillovers across Mexico's states. *The Journal of International Trade & Economic Development*, 0(iFirst), 1-20.
- Freeman, R. B. (2009). *Labor Regulations, Unions, and Social Protection in Developing Countries: Market distortions or Efficient Institutions?* (NBER Working Papers No. 14789). National Bureau of Economic Research, Inc.
- Hsieh, C.-T., & Klenow, P. (2012). The Life Cycle of Plants in India and Mexico. *NBER Working Paper*(18133).
- ILO. (2003). Guidelines concerning a statistical definition of informal employment. In *17th international conference of labour statisticians*. Report of the Conference Doc. ICLS/17/2003/R, Geneva: International Labour Office.
- Khamis, M. (2012). A Note On Informality In The Labour Market. *Journal of International Development*, 24(7), 894-908.

- Levy, S. (2010). *Good intentions, bad outcomes: Social policy, informality, and economic growth in Mexico*. Brookings Institution Press.
- Loayza, N. V., Servén, L., & Sugawara, N. (2009). *Informality in Latin America and the Caribbean* (Policy Research Working Paper Series No. 4888). The World Bank.
- Maloney, W. F. (2004). Informality Revisited. *World Development*, 32(7), 1159-1178.
- Markusen, J. R., & Venables, A. J. (1999). Foreign direct investment as a catalyst for industrial development. *European Economic Review*, 43(2), 335-356.
- McMillan, M., & Rodrik, D. (2012). *Globalization, structural change, and productivity growth* (IFPRI discussion papers No. 1160). International Food Policy Research Institute.
- Melitz, M. J. (2003). The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity. *Econometrica*, 71(6), 1695-1725.
- OECD. (2009). *Is Informal Normal? Towards More and Better Jobs in Developing Countries*. OECD Publishing, Paris.
- Roodman, D. (2009). How to do xtabond2: An introduction to difference and system GMM in Stata. *Stata Journal*, 9(1), 86-136.
- Schneider, F., Buehn, A., & Montenegro, C. E. (2010). *Shadow economies all over the world: new estimates for 162 countries from 1999 to 2007* (Policy Research Working Paper Series No. 5356). The World Bank.
- Stel, A., Storey, D., & Thurik, A. (2007). The Effect of Business Regulations on Nascent and Young Business Entrepreneurship. *Small Business Economics*, 28(2), 171-186.
- Windmeijer, F. (2005). A finite sample correction for the variance of linear efficient two-step GMM estimators. *Journal of Econometrics*, 126(1), 25-51.
- Wooldridge, J. (2002). *Econometric Analysis Cross Section Panel*. Massachusetts Institute of Technology.