

ORIGINAL ARTICLE

Open Access



Wage inequality and informality: evidence from Mexico

Chiara Binelli^{1,2}

Correspondence:

chiara.binelli@gmail.com

¹Department of Economics,
University of Southampton,
Southampton, UK

²Center for Research and Social
Progress, Piacenza, Italy

Abstract

While several studies have documented the expansion of the informal sector and its detrimental impact on development, few have noted that informality and wage inequality tend to move together. Using Mexico as a case study, I show that between 1987 and 2002 wage inequality within informal workers accounted for over 60 % of total wage inequality and that the Mexican financial crisis of the mid-1990s increased the share of informal workers and, via this, wage inequality. The results provide supportive evidence that in Mexico higher wage dispersion is one of the channels through which informality negatively affects development.

JEL codes: O17, J31, C36.

Keywords: Wage inequality, Informality, Mexico

1 Introduction

Informal labour markets are widespread, particularly in low- and middle-income countries. In their survey of the literature, Schneider and Enste (2000) find that at the beginning of the 1990s in most countries in Africa, Asia and Central and South America the informal sector accounted for over half of aggregate GDP. A leading example of the prevalence of informal markets is Latin America, where by the end of the 1990s the size of the informal sector accounted for 40–50 % of the labour force (Maloney 1999; Attanasio et al. 2004).

The expansion of the informal sector is a response to weak institutions and a burdensome regulatory environment. A growing literature has shown that countries with poorly functioning institutions, labour rigidities, a heavy tax burden and high levels of corruption tend to have large informal sectors (Loayza 1996; Johnson et al. 1998; Botero et al. 2004; Vuletin 2009) and lower aggregate welfare (Meghir et al. 2015). Therefore, unsurprisingly, the bigger the informal sector, the lower the level of economic development. Pratap and Quintin (2006) find that informal output represents 10–15 % of GDP in most high-income countries, compared to 25–80 % in most low- and middle-income countries.

While several studies have documented the expansion of the informal sector, few have noted that changes in the size of this sector tend to move together with changes in wage inequality. An example of this trend is Latin America, where in the 1990s most countries experienced significant changes both in informality and in wage inequality. Among Latin American countries, Mexico provides an interesting setting to study the relationship between wage inequality and informality since the co-movement between changes in the

share of the informal workers and in wage inequality was particularly evident. Binelli and Attanasio (2010) present a descriptive analysis showing how in the 1990s changes in informality tracked closely changes in wage inequality: in the first half of the 1990s, both the proportion of workers employed in the informal sector and wage inequality increased, reached a peak in 1995 and decreased afterwards.

The descriptive evidence presented by Binelli and Attanasio (2010) is interesting but is neither complete nor informative of the direction of causation. First, how did the changes in informality contribute to the changes in wage inequality? In particular, how much of the change in overall wage inequality is due to changes in wage inequality within formal and informal workers relative to changes in wage inequality between formal and informal workers? Second, what is the mechanism through which changes in informality affected changes in wage inequality? In the mid-1990s, Mexico was hit by a severe financial crisis, the peso crisis, which resulted in substantial income and consumption losses (McKenzie 2001), a massive devaluation of the national currency (Whitt 1996) and an increase in wage inequality (Verhoogen 2008). Was the increase in informality one of the channels through which the peso crisis affected wage inequality? In other words, did the peso crisis affect informality and, via this, wage inequality?

In order to answer these questions, I first compute a wage inequality decomposition to quantify the change in overall wage inequality that is due to changes in wage inequality within formal and informal workers relative to changes in wage inequality between formal and informal workers, and I then assess the impact of the peso crisis on the share of informal workers and, via this, on wage inequality. In particular, I propose a simple mechanism of a direct effect of informality on wage inequality by using the peso crisis as an instrumental variable. The peso crisis of the mid-1990s acted as an exogenous macroeconomic shock that hit the Mexican economy and resulted in a significant rise in unemployment among formal workers. Mexico has no programme of unemployment compensation, so only those with some economic comfort can afford to be unemployed (Martin 2000). For the majority of the jobless workers, the informal sector became the only viable option, and, consistently, the data show that the proportion of workers employed in the informal sector sizably increased. While in the formal sector wages are kept stable by collective bargaining and employment protection legislation, in the informal sector wages are lower and unregulated, thus more flexible by design. An increase in the proportion of informal workers therefore increased aggregate wage inequality.

Both the results of the wage inequality decomposition and the instrumental variable regressions provide supportive evidence that in Mexico the increase in informality resulted in higher wage inequality. The empirical analysis provides a research strategy and framework to assess the impact of informality on wage inequality, which could be tested in other countries to assess the robustness of the findings in different settings and labour market institutions.

The remainder of the paper is organized as follows. Section 2 describes the data; Section 3 performs a within-between formal and informal workers' decomposition of the Theil index of wage inequality; Section 4 reviews the previous literature on informality and wage inequality, and discusses how the peso crisis affected informality and, via this, wage inequality. Section 5 presents the empirical strategy, the main results and a detailed discussion of robustness checks. Section 6 briefly concludes. The Appendix includes all tables and figures.

2 Data

I use micro data from the Mexican Employment Survey (*Encuesta Nacional de Empleo Urbano* or ENEU) from 1987–2002. The ENEU is a quarterly household survey with a rotating panel structure similar to the US Current Population Survey: households are interviewed for five consecutive quarters, and in each quarter, 20% of the households are replaced by new households that are interviewed for the first time. The survey is conducted annually by the Mexican National Statistical Institute (*Instituto Nacional de Estadística, Geografía e Informática* or INEGI) and covers only urban areas, with municipalities being the primary sampling units. The sampling scheme has changed over time as a number of smaller municipalities have progressively entered the sample.

The ENEU is the only Mexican household survey continuously available since the late 1980s that collects detailed labour market information and a large array of socioeconomic characteristics. As such, it has been widely used for studies of the Mexican labour market, including several studies on changes in the wage distribution (e.g. Verhoogen 2008; Binelli and Attanasio 2010; Bosch and Manacorda 2010; Binelli 2015). For all individuals at least 12 years old, the survey contains detailed employment information with several questions on occupational status, type and characteristics of employment, characteristics and sector of main and secondary job, contract type, number of hours worked, monthly wages, unemployment status and duration.

Importantly for the empirical analysis conducted in this paper, the ENEU accurately reports earnings and social security contributions paid by the worker's employer both in the private and in the public sector, which I use to identify the workers employed in the formal and in the informal sector. In particular, I define as informal a worker who does not pay social security contributions in either the private or in the public sector. The definition of earnings in the survey refers to monthly earnings received from the main job net of all labour taxes and social security contributions. For those paid by the week, the survey transforms weekly earnings into monthly earnings by multiplying the former by 4.3. Similar adjustments are used for workers paid by the day or every 2 weeks. Hourly wages are computed as the ratio between monthly earnings and hours worked in the main occupation.

The sample selection criteria follow Binelli and Attanasio (2010): I consider all adults aged between 25 and 60 that are actively working at the time of the interview in all municipalities included in the fourth quarter of each survey year between 1987 and 2002. I adjust wage data for inflation by using the Mexican national CPI of June 2002.

I consider four types of workers: self-employed (individuals working on their own with no employees at their dependency), employers (individuals working on their own with some employees at their dependency), wage and piece workers, and I define four skill groups according to the highest level of completed education: less than primary (0), primary (1), secondary (2) and post-secondary (3) education. The average skill level and age in the sample are, respectively, 1.1 and 40, the sample is predominantly male and around 87% of the individuals are working. Of those working, over 60% are employees, around 22% are self-employed and only around 8 and 7% are working, respectively, as employers and piece workers. Overall, more than half of the workers in the sample have a job in the informal sector.

As in Binelli and Attanasio (2010), Fig. 1 plots the proportion of informal workers among the 25–60 working age population and the formality premium, which is computed

as the average wage differential between formal and informal workers. Confirming the evidence presented in a number of previous studies (Maloney 2004; Bosch and Maloney 2006), Fig. 1 shows that in the 1990s the share of informal workers accounted for around half of the urban workforce in Mexico. Further, between 1994 and 1995, the proportion of informal workers jumped up by 4 percentage points to reach the value of 56 % in one year. In 1998, it was back to the pre-1995 level, and it stabilized around a value of 52 % thereafter. The formality premium increased in the first half of the 1990s and decreased substantially after 1998.

Binelli and Attanasio (2010) present a detailed descriptive analysis of informal employment in Mexico between 1987 and 2002 and distinguish between four types of informal jobs: self-employed (individuals working on their own with no employees at their dependency), employers (individuals working on their own with some employees at their dependency), wage and piece workers. They find that in each year of the sample self-employment represents the largest share of the informal sector accounting for over 40 % of the informal labour force, while informal salaried workers represent between 30 and 35 % of informal workers. In addition, the proportion of piece workers is stable at around 10 % throughout the sample period. The evidence in Binelli and Attanasio (2010) show that the increase in informal employment at the time of the peso crisis was due to an increase in informality among wage workers.

2.1 Wage inequality

The four panels a, b, c and d presented in Fig. 2 reproduce Figure 7 in Binelli and Attanasio (2010) and describe the changes in wage inequality between 1987 and 2002 using the four most commonly used measures of aggregate wage dispersion: the variance of the logarithm, the Gini coefficient, the 90th–50th and the 50th–10th percentile ratios of hourly real wages. In order to assess the sensitivity of the results to the use of the sample that excludes the zeros, graphs b, c and d also report the results obtained when the zeros are included.

As documented by an extensive previous literature on Mexico (for example, Hanson and Harrison 1999; Robertson 2004; Bosch and Manacorda 2010), all inequality indicators increased sharply between the end of the 1980s and 1995 and declined afterwards. Between 1987 and 1995, the variance of the logarithm and the Gini coefficient for real hourly wages increased, respectively, by 31 % and by 5 percentage points. Between 1995 and 2002, the variance of the logarithm fell by around 20 % and the Gini coefficient decreased by around 7 percentage points. As expected, in each year of the sample, wage inequality is higher when we use the sample that includes the zeros. However, the trends remain the same: wage inequality increased in the first half of the 1990s and decreased afterwards.

3 Wage inequality decomposition

In order to quantify the change in overall wage inequality that is due to changes in wage inequality within formal and informal workers relative to changes in wage inequality between formal and informal workers, I compute a Theil wage inequality decomposition. I use the Theil index since it is the most commonly used indicator in the class of inequality indexes that are decomposable without a residual term (Bellù and Liberati 2006).¹

Let us consider the wage distribution of a given population of workers n in year t and let us divide the n population into two main groups (m): formal ($m = f$) and informal ($m = I$) workers. Let us denote the population mean wage in year t as \bar{w}^t , the mean wage of workers in group k as \bar{w}_k^t and the population share of workers in group k in year t as $\frac{n_k^t}{n^t}$. In each year t , we can decompose total wage inequality in the between-group wage differential between formal and informal workers and the within-group wage differential:

$$T = \underbrace{\sum_{k=1}^m \left(\frac{n_k}{n} \frac{\bar{w}_k}{\bar{w}} \right) T_k}_{\text{WITHIN}} + \underbrace{\sum_{k=1}^m \frac{n_k}{n} \left(\frac{\bar{w}_k}{\bar{w}} \right) \ln \left(\frac{\bar{w}_k}{\bar{w}} \right)}_{\text{BETWEEN}}, \quad (1)$$

where T denotes the Theil index for overall wage inequality in the n population of workers in a given year t , which, for simplicity, is omitted.

The within part of the decomposition is the weighted average of the Theil inequality indexes of each k group (T_k) with weights given by the total wage share of group k (the product of group k 's population share and relative mean wage). I compute the share of formal and informal workers in total wages, and I calculate the Theil index of each group separately. This index is then multiplied by the population share in order to identify the contribution of each group to the within inequality. The sum of all these contributions gives the within component of the wage decomposition.

The between part of the decomposition is the Theil index, calculated using subgroup mean wages \bar{w}_k instead of actual incomes. The computation of the between component follows the logic of replacing actual wages in each group with the average wage of the group. In particular, I calculate mean wages separately for formal and for informal workers, and I then replace each actual individual wage with the mean wage of the group the individual belongs to. The between term then computes the Theil index of this fictitious wage distribution. Since wages are all equal and they only differ between groups, the Theil of this fictitious distribution captures the dispersion of wages attributable to the difference between formal and informal workers, that is, the between-group component of total wage inequality.

Once the within and the between components have been computed, the decomposability of the Theil index allows to check the exactness of the decomposition by verifying that in each year the sum of the within and the between component is equal to the total wage inequality computed by the Theil for the overall wage distribution.

Tables 1 and 2 present the results of the Theil decomposition for each year of the sample. While Table 1 confirms the results of Binelli and Attanasio (2010) that wage inequality is much higher among informal than among formal workers, Table 2 additionally shows that in each year the within component accounts for almost all wage inequality. In order to see this clearly, we can use the wage decomposition in Eq. (2) to directly compute the relative contribution of changes in wage inequality within informal workers. We can do this by computing the product of the share of informal workers in total wages and the Theil index for informal workers as well as by taking the ratio of this product and the Theil for the overall wage distribution. By doing so, I find that in each year of the sample, wage inequality within informal workers accounts for more than 60 % of overall wage inequality, with peaks of up to 70 % in 1995, 1996 and 1997.² As a consequence, an increase in the

share of informal workers, and thus an increase in the weight of the within component ($\frac{w_i}{n}$), maps directly into an increase in overall wage inequality.

4 Linking informality and wage inequality

While a vast literature has documented and studied wage inequality and informality in isolation, very few papers have linked the two either by developing theoretical models or by performing empirical analysis.

4.1 Previous literature

On the theory side, Chong and Gradstein (2007) develop a model in which an increase in income inequality causes a bigger informal sector by lowering the relative benefits from becoming formal, with the effect being stronger when the institutions and the protection of property rights are weak in the formal sector. They find empirical support for a positive impact of inequality on the size of the informal sector using cross-country regressions over the period 1990–2000.

Other papers have proposed theoretical models that study the reverse relationship of informality on inequality as a side effect of changes in tax revenues and in the availability of public goods. Johnson et al. (1998) develop a model of the public sector where the share of the informal sector is negatively related to tax revenues and, thus, to the availability of public goods. Loayza (1996) also models the supply of congestible public goods as a negative function of the share of the informal sector and finds supporting empirical evidence among Latin American economies.

Related to this literature, Goldberg and Pavcnik (2003) develop a model to establish a link between informality and trade liberalization. In their model, trade opening leads to an increase in the size of the informal labour market via a simple mechanism: trade reforms expose formal establishments to increased foreign competition to which firms in the formal sector respond by laying off workers who subsequently seek employment in the informal sector. Goldberg and Pavcnik (2003) and Attanasio et al. (2004) use data for Colombia and Brazil and find that opening to trade increased the size of the informal sector only in Colombia, by a small amount, and only in the sectors where the tariff cuts were the largest. In contrast, Aleman-Castilla (2006) analyses the relationship between trade liberalization and informality using a heterogeneous firm's model that predicts a negative effect of trade on informality by making it more profitable for some firms to enter the formal sector by forcing the least productive informal firms to exit the industry and by inducing the most productive formal firms to engage in trade, finding some supporting evidence for Mexico.

On the empirical side, to the best of my knowledge, the only three papers that study the direct impact of informality on wage inequality are Rosser et al. (2000 and 2003) and Binelli and Attanasio (2010). Rosser et al. (2003) estimate a cross-country regression for 18 transition economies in the 1990s and find that the share of the informal sector has a positive and significant impact on the level of the Gini index for wage inequality. Binelli and Attanasio (2010) use time-series data for Mexico between 1987 and 2002 and find that an increase in informality is positively and statistically significantly correlated with a substantial increase in wage inequality.

Despite providing interesting results, the empirical analyses of both Rosser et al. (2000 and 2003) and Binelli and Attanasio (2010) suffer from endogeneity: without a valid

instrument to act as an exogenous source of variation, the direction of causation between informality and wage inequality cannot be established, and the analyses remain of a pure descriptive nature.

4.2 Peso crisis and informality

I propose a simple mechanism to establish a direct effect of informality on wage inequality. I argue that the Mexican financial crisis of the mid-1990s can be used as an instrument to identify the impact of informality on wage inequality. In order to use of the peso crisis as an instrumental variable, I will assume that the crisis affected wage inequality only through informality. I will test the robustness of this assumption in Section 5.2.2.

In January 1994 Mexico became a member of the Organization for Economic Cooperation and Development (OECD) and entered the North American Free Trade Agreement (NAFTA) with the US and Canada. In December 1994, following a drain on its foreign exchange reserves, the peso was devalued by 40 % (Whitt 1996). Between 1994 and 1996, the Mexican GDP decreased by around 7 %. The international response to the crisis was immediate, and assistance was promptly provided by the US and the IMF. As part of the rescue package, in March 1995, the Mexican government released a new economic plan to address the economic requirements set by the US and the IMF. The recovery was rather quick, and by the end of 1995, Mexico had reentered the international capital markets. The economic and financial recovery corresponded to a decrease in the share of informal workers, which by 1998 was back to the pre-crisis level.

A large literature has studied the impact of the peso crisis on a number of outcomes, including household income and consumption, labour supply, wage inequality and informality. Using individual-level data from the Mexican Consumption and Expenditure Survey (ENIGH), McKenzie (2001) finds that the peso crisis had no significant effect on labour supply and caused income and consumption to drop for all groups, although the relative impact of the crisis differed by education and residence of the household head in urban and rural areas. In particular, McKenzie (2001) computes net income from all sources excluding income from capital transactions and finds that the more highly educated urban workers had larger falls in income than less educated and rural workers.

Lopez-Acevedo and Salinas (2000) find that income inequality in Mexico improved during the financial crisis. They attribute this decrease in income inequality to the crisis having had a greater impact on the labour earnings of workers in non-tradeable sectors, such as financial services, and on individuals in the top income deciles. More recently, Verhoogen (2008) finds that the peso crisis induced a quality upgrading mechanism such that more productive plants increased exports, quality and wages relative to less productive plants within the same industry, thus increasing within-industry wage inequality.

To the best of my knowledge, Martin (2000) is the only paper that studies the direct impact of the peso crisis on informality. Consistently with the evidence presented in Fig. 1, Martin (2000) finds that during the peso crisis the informal sector grew. In particular, both the proportion of workers employed in enterprises with five or fewer employees and the proportion of workers who do not receive any employment benefits increased.

The peso crisis acted as an exogenous macroeconomic shock that hit the Mexican economy and resulted in a significant rise in unemployment among formal workers. Mexico has no programme of unemployment compensation, which, as Martin (2000) points out,

means that only those with some economic comfort can afford to be unemployed. For the jobless workers, the informal sector became the only viable option; consistently, the data show that the share of informal workers as a proportion of the labour force sizably increased (Binelli and Attanasio 2010; Martin 2000). While in the formal sector wages tend to be kept stable by collective bargaining agreements and employment protection legislation, in the informal sector, wages are, on average, lower than in the formal sector and unregulated and therefore more flexible by design. Consistently, for all indicators and both at the top and at the bottom of the wage distribution, inequality is much higher among informal workers (Binelli and Attanasio 2010), and the formality premium is high and substantial throughout the 1990s (Fig. 1).

As the financial crisis hit the economy, the increase in the proportion of informal workers translated into an increase in the proportion of wages that are lower and with a relatively higher variance, thus increasing overall wage dispersion. As shown by the wage decomposition in Eq. (2), an increase in the share of informal workers, and thus an increase in the weight of the within component ($\frac{m}{n}$), maps directly into an increase in overall wage inequality. Therefore, wage inequality increased because of an increase in the proportion of informal workers due to the worsening labour market opportunities caused by the financial crisis. Consistently with this mechanism, as discussed in Section 3, in the years of the peso crisis, inequality within informal workers accounts accounted for up to 70 % of overall wage inequality.

5 Empirical analysis

In order to assess the impact of the peso crisis on informality and, consequently, on wage inequality, I specify a system of equations where wage inequality is a function of informality and informality is affected by the peso crisis. In particular, following Binelli and Attanasio (2010), I specify a first equation in which wage inequality is a function of the share of informal workers, education and age shares to control for changes in skill and age composition, and a linear time trend to control for year effects,³ and I then specify a second equation where the share of informal workers is a function of the peso crisis and education and age shares. Therefore, I specify the following model:

$$\text{Inequal}_t = \gamma_0 + \gamma_1 I_t + \gamma_2 \text{year}_t + \gamma_3 \text{skillcomp}_t + \gamma_4 \text{agecomp}_t + \varepsilon_t \quad (2)$$

$$I_t = \alpha_0 + \alpha_1 \text{Crisis}_t + \alpha_2 \text{year}_t + \alpha_3 \text{skillcomp}_t + \alpha_4 \text{agecomp}_t + u_t \quad (3)$$

where Inequal_t is an aggregate measure of wage inequality in year t , which is either the Gini index or the variance of log hourly real wages in year t . I_t is the share of informal workers in year t . skillcomp_t is a vector of education shares, that is, $\text{skillcomp}_t = \{\text{share } sk2_t, \text{share } sk3_t, \text{share } sk4_t\}$, where $\text{share } skj_t$ is the proportion of those in the j th education group in year t . agecomp_t is a vector of population shares by age group, that is, $\text{agecomp}_t = \{\text{share } age2_t, \text{share } age3_t\}$, where $\text{share } agej_t$ is the proportion of those in the j th age group in year t . year_t is a linear time trend to control for any time-varying aggregate trend that could affect wage inequality. Crisis_t is a period dummy that is equal to one for the years of the peso crisis between 1994 and 1996, and zero otherwise.

I compute averages of the variables of interest for each year between 1987 and 2002, and I estimate the model using a three-stage least generalized least squares estimator (3SLS). The 3SLS estimator allows producing consistent estimates of the model's parameters, and the generalized least squares account for the correlation structure in the error

terms across Eqs. (2) and (3) (Davidson and MacKinnon 1993, pp. 651–661, and Greene 2012, pp. 331–334).⁴

5.1 Main results

Tables 3 and 4 present the main results. Confirming the direction and sign of the impacts discussed in Section 4.2, the results show that the peso crisis has a statistically significant and positive impact on informality and that informality has a strong and positive impact on wage inequality. The magnitude of the impact of informality on wage inequality is impressively consistent across the two wage inequality measures, and the positive sign of the impact confirms the results of the previous literature (Rosser et al. 2003; Binelli and Attanasio 2010).

It is interesting to compare the instrumental variable (IV) results reported in Tables 3 and 4 with the results obtained when estimating Eq. (2) using ordinary least squares (OLS). Unreported results, all available from the author upon request, confirm the findings of Binelli and Attanasio (2010) of a positive and statistically significant impact of informality on wage inequality. However, as expected, the size of the impact differs. In particular, compared to the estimated effects using OLS, the IV impact of informality on wage inequality is, respectively, 30 % smaller when the dependent variable is the Gini index, and 1.2 times smaller when the dependent variable is the variance of log hourly real wages, which is evidence of a substantial upward bias in the OLS estimates.⁵

5.2 Robustness

By estimating the system of Eqs. (2) and (3), I treat informality as an endogenous variable, and I use the peso crisis as an instrumental variable that affects informality and, via this, wage inequality. In addition to this, the generalized least squares account for the correlation structure in the disturbances across the two equations.

The validity of any instrumental variable strategy depends on two main conditions: first, the presence of a strong correlation between the instrument and the endogenous variable and, second, the absence of any correlation between the instrument and the error term (exogeneity condition). The results in Table 4 provide a direct test of the validity of the first condition. On the contrary, the validity of the exogeneity condition cannot be tested directly; the robustness checks in this Section provide an indirect test of its validity.

5.2.1 Linear instrument

As discussed in Section 4.2, the impact of the peso crisis on the Mexican economy continued until 1996; however, international assistance was promptly provided, and after 12 months of economic turmoil, in December 1995, Mexico started a fast recovery. Consistently, both wage inequality and the share of informal workers reached a peak in 1995 and then started to decrease returning to the pre-crisis levels by 1998.

As discussed by Binelli and Attanasio (2010), even if both wage inequality and the share of informal workers increased between 1994 and 1996, these overall changes were split into an initial increase between 1994 and 1995 and a decrease between 1995 and 1996. Given these changes, a concern is that the peso crisis defined with a dummy variable that is equal to one for each year between 1994 and 1996 makes a linearity assumption which does not take into account the ups and downs in the data series. I address this concern by constructing a dummy variable that is equal to one for the years 1994 and 1995 (when both

series increased) and zero otherwise and accordingly re-estimate the system of Eqs. (2) and (3).

Tables 5 and 6 present the results. Consistently with the main results, the peso crisis has a strong positive impact on informality, and the size of the correlation between informality and wage inequality is statistically significant, positive, and of even a larger size than in the main results presented in Tables 3 and 4.

5.2.2 Computing bounds

In order to use of the peso crisis as an instrumental variable, I have to assume that the crisis affected wage inequality only through informality. The exogeneity assumption of no correlation between the instrument and the error term is not directly testable, and it is violated when there are variables that are potentially important to explain the outcome and are correlated with the endogenous variable but are not included in the analysis due, for example, to lack of data.

When considering the system of Eqs. (2) and (3), it is easy to think of a number of variables, such as unemployment and poverty rates, that are likely to be important determinants of wage inequality and could be correlated with the share of informal workers but are not included in the analysis. The choice of excluding these variables is due to data limitation: since I only have data between 1987 and 2002, and therefore only 16 years of data, including additional variables would quickly saturate the model. In addition to omitted variable bias, Eq. (2) could suffer from reverse causality since inequality could affect informality (see Section 4.1).

Therefore, the validity of the exogeneity condition is a serious concern. In order to address this concern, I use the method developed by Nevo and Rosen (2012) to allow for correlation between the instrument and the error term, and estimate bounds for the impact of informality on wage inequality. As in other recent applications of this approach (e.g. Binelli and Rubio-Codina 2013; Fitzsimons and Malde 2014), the bounds are used to investigate the extent to which relaxing the exogeneity assumption affects the main empirical findings.

Following Nevo and Rosen (2012), I assume that:

1. *The correlation between the endogenous variable and the error term and between the instrument and the error term has the same sign.* An extensive literature has shown that the informal sector in Mexico attracts the least skilled and the youngest workers (e.g. Maloney 1999; Binelli and Attanasio 2010).⁶ Therefore, the most likely correlation between the share of informal workers and the error term in Eq. (2) is positive. Confirming a likely positive correlation between the share of informal workers and the error term in Eq. (2), an OLS estimation of Eq. (2) returns an impact of informality on wage inequality that is larger in magnitude than the impact estimated using IV.⁷ At the same time, the potential correlation between the financial crisis and the error term in Eq. (2) is likely to be positive since the financial crisis has been identified as a determinant of an increase in within-industry wage inequality (Verhoogen 2008); therefore, if the financial crisis had any residual direct effect on wage inequality, this effect would be positive.
2. *The correlation between the instrument and the error term is less strong in absolute terms than the correlation between the endogenous variable and the error term.* This assumption weakens the exogeneity condition by allowing some correlation between

the instrument and the error term, and its validity rests on a relative comparison of strength of correlations that is not directly testable. The only admissible comparison between these two correlations is a comparison that is theory based.

The share of informal workers is clearly endogenous. Workers self-select into the formal or informal sector, and this choice affects wage inequality since wage dispersion differs in the two sectors (see Section 3). Therefore, informality and the error term in Eq. (2) are correlated, that is, $\text{Cov}(I_t, \varepsilon_t) \neq 0$. Computing bounds also allows for correlation between the instrument and the error term, that is, $\text{Cov}(\text{Crisis}_t, \varepsilon_t) \neq 0$. Assumption 2 is satisfied when $|\text{Cov}(\text{Crisis}_t, \varepsilon_t)| < |\text{Cov}(I_t, \varepsilon_t)|$. The peso crisis acted as an aggregate exogenous shock that hit the Mexican economy and informality directly (Martin 2000 and the discussion in Section 4.2). Given Eq. (2), $\text{Cov}(\text{Crisis}_t, \varepsilon_t)$ could be different from zero if there were other channels different from informality through which the financial crisis affected wage inequality, once having accounted for age and skill composition and time-varying aggregate trends. As already discussed, McKenzie (2001) finds that the peso crisis had no significant effect on labour supply and caused income and consumption to drop for all groups, although the relative impact of the crisis differed by education and residence of the household head in urban and rural areas. In particular, he finds that the more highly educated, urban workers had larger falls in income than less educated and rural workers. Therefore, urbanity and education are two important variables to control for when estimating the impact of the financial crisis on informality and, via this, on wage inequality. I estimate the system of Eqs. (2) and (3) using only urban data, and Eq. (3) includes a vector of education shares to control for skill composition.

More recently, Verhoogen (2008) finds that the peso crisis induced a quality upgrading mechanism such that more productive plants increased exports, quality and wages relative to less productive plants within the same industry, thus increasing within-industry wage inequality. While, unfortunately, the aggregate analysis in Eqs. (2) and (3) does not allow to control for within-industry wage inequality directly, I can re-estimate Eqs. (2) and (3) by restricting the sample to the workers that are not employed in an industry, so that wage inequality, informality and the rest of the aggregate variables in Eqs. (2) and (3) do not reflect any within-industry change. As discussed in Section 2, I consider four types of workers: self-employed, employers, wage and piece workers. While employers, wage and piece workers could be employed in an industry, self-employed are individuals working on their own with no employees at their dependency. In the sample, there are 247,672 individuals that are self-employed. By using the subsample of self-employed to compute the aggregate variables and re-estimate Eqs. (2) and (3), I obtain substantively similar results to the ones presented in Tables 3 and 4, that is, the peso crisis positively affects informality, which has a positive impact on wage inequality.

Having accounted for the factors that have been identified as other potential channels through which the financial crisis affected wage inequality, it is reasonable to assume that any residual potential correlation between the instrument and the error term induced by omitting these variables would be smaller in size than the correlation between the endogenous variable and the error term.

3. *The correlation between the instrument and the endogenous variable is negative, and the correlation between the endogenous variable and the error term is positive. As*

already discussed, the most likely correlation between the share of informal workers and the error term in Eq. (2) as well as the correlation between the share of informal workers and the peso crisis is positive. Following Fitzsimons and Malde (2014), we can computationally satisfy the assumption that the correlation between the instrument and the endogenous variable is negative by defining the instrument as $-\text{Crisis}$ and accordingly multiply the obtained bounds by minus.

If assumptions 1, 2 and 3 are satisfied, Nevo and Rosen (2012) show that it is possible to obtain a two-sided bound given by

$$B^* = [\beta^{IV}, \beta^{IV^*}]$$

where β^{IV^*} is the upper bound and is an IV estimator in which the instrument is defined as $\text{Crisis}^* = (\sigma_{\text{Crisis}}I - \sigma_I \text{Crisis})$ where σ_I and σ_{Crisis} are, respectively, the standard deviation of the endogenous variable and of the instrument. In other words, the upper bound, β^{IV^*} , is given by the IV estimate obtained when using the transformed instrument, Crisis^* , which is exogenous by construction. The lower bound, β^{IV} , is given by the IV estimate obtained when using the (imperfect) instrument Crisis . Nevo and Rosen (2012) show that the larger the correlation between the instrument and the endogenous regressor, the tighter the upper bound.

I estimate the upper bound β^{IV^*} of the effect of informality on wage inequality for the main results reported for models 1 and 2 in Table 2. I find that $\beta^{IV^*} = 0.893$ when the dependent variable is the Gini index, so that for model 1 $B = [0.773, 0.893]$, and that $\beta^{IV^*} = 0.798$ when the dependent variable is the variance of log wages, thus in model 2 $B = [0.738, 0.798]$. Therefore, when the exogeneity assumption is relaxed, the IV effect of informality on wage inequality is bounded tightly, and the size of the upper bounds confirms a strong positive effect of informality on wage inequality.

6 Conclusions

In several countries changes in the size of the informal sector and in aggregate wage inequality tend to move together. An interesting case study is provided by Mexico, where in the 1990s the co-movement between informality and wage inequality was particularly strong. Using Mexican data from 1987–2002, I first show that in each year of the sample wage inequality within informal workers accounted for over 60% of total wage inequality, and I then use an instrumental variable strategy to show that the peso crisis of the mid-1990s was one of the channels through which informality affected wage inequality. In particular, I show that at the onset of the financial crisis the proportion of workers employed in the informal sector increased, while both wage inequality and the proportion of informal workers decreased to the pre-crisis levels once the economy started to recover and the crisis resolved. Verhoogen (2008) finds that the Mexican peso crisis contributed to higher wage inequality by increasing within-industry wage dispersion. This paper shows that the expansion of the informal sector was an additional channel through which the Mexican peso crisis increased wage inequality.

While the validity of the results is limited to Mexico, the findings provide an interesting research hypothesis which could be tested in different country contexts and has important policy implications. In particular, the increase in wage inequality in the decade of the 1990s in Mexico as well as in several other countries has been associated with a number

of negative outcomes, such as low investments in higher education due to non-linearities in the returns to schooling (Manacorda et al. 2010; Binelli 2015). Likewise, informality has been identified as detrimental to economic development (Pratap and Quintin 2006). The results of this paper suggest that higher wage dispersion is one of the channels through which informality negatively affects development.

Endnotes

¹The Theil index satisfies the four basic desirable properties of an inequality measure (Shorrocks 1980); as such, it has been extensively used in several inequality analyses (Galbraith 2012).

²All results are available from the author upon request.

³As discussed in Section 2, I define four education groups according to the highest completed schooling level: less than primary, primary, secondary and post-secondary. Following Binelli and Attanasio (2010), I define three age groups: the young group of those aged between 25 and 35, the middle group of those aged between 36 and 50, and the old group of those aged between 51 and 60.

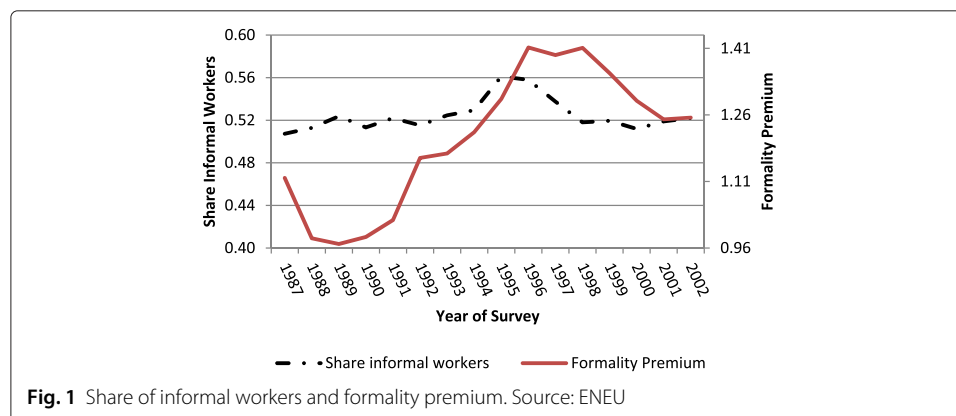
⁴3SLS produces instrumental variable estimates from a three-step process. The main difference between the standard two-stage least squares (2SLS) with robust standard errors and the 3SLS method lies in the assumptions on the correlation of the error terms between the first- and second-stage equation, and therefore in the computation of the standard errors. The results when estimating the model using 2SLS with robust standard errors return identical point estimates and smaller standard errors in the second stage. All 2SLS results are available from the author upon request.

⁵Confirming the robustness of the relationship between informality and wage inequality, both an OLS regression of the model estimated without the time trend and an OLS regression of wage inequality on informality only, return a positive and statistically significant impact of informality on wage inequality. All results are available from the author upon request.

⁶Confirming the results of the previous literature, by regressing an indicator variable that equals one if an individual works in the informal sector on a number of observables (age, age squared, education level, sector of employment, dummies for year of survey and municipality of residence), I find that informal workers are predominantly young, low-educated and working as employers or self-employees. All results are available from the author upon request.

⁷All results are available from the author upon request.

Appendix



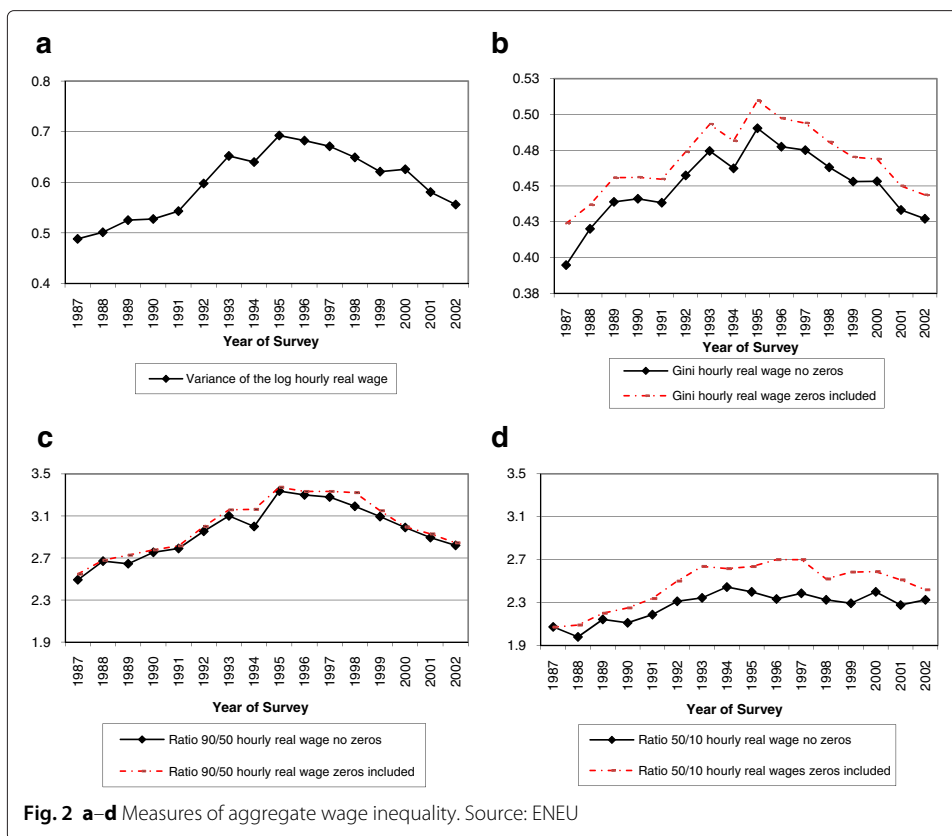


Table 1 Wage inequality among formal and informal workers by year using Theil inequality index

Year	Formal workers	Informal workers	% difference
1987	0.169	0.372	120
1988	0.241	0.484	101
1989	0.333	0.529	59
1990	0.312	0.533	71
1991	0.338	0.480	42
1992	0.319	0.601	89
1993	0.355	0.617	74
1994	0.341	0.530	55
1995	0.393	0.718	83
1996	0.329	0.576	75
1997	0.327	0.594	81
1998	0.332	0.539	62
1999	0.304	0.528	74
2000	0.301	0.543	81
2001	0.294	0.452	54
2002	0.307	0.450	47

Table 2 Wage decomposition formal-informal workers by year using Theil inequality index

Year	Within component	Between component
1987	0.272	0.000
1988	0.366	0.006
1989	0.436	0.005
1990	0.425	0.000
1991	0.412	0.002
1992	0.464	0.001
1993	0.493	0.000
1994	0.441	0.001
1995	0.576	0.001
1996	0.466	0.002
1997	0.470	0.002
1998	0.440	0.003
1999	0.420	0.002
2000	0.425	0.001
2001	0.376	0.001
2002	0.382	0.002

Table 3 Regression of wage inequality on the share of informal workers in model 1 and in model 2. System of equations with peso crisis defined as a dummy = 1 for 1994, 1995 and 1996

	Model 1	Model 2
Informality	0.773 (0.244)	0.738 (0.263)
Share skill 2	0.985 (0.904)	-0.359 (0.362)
Share skill 3	1.878 (0.621)	0.458 (0.175)
Share skill 4	0.001 (0.013)	-0.003 (0.010)
Share age 2	-1.361 (0.909)	-0.631 (0.737)
Share age 3	-1.151 (0.508)	-0.619 (0.135)
Year	-0.005 (0.007)	0.002 (0.012)
Constant	10.659 (12.778)	-4.617 (26.076)

Standard errors in parenthesis. Sample of workers aged 25–60. Dependent variable: Gini (model 1); var log wages (model 2)

Table 4 Regression of the share of informal workers on the peso crisis for model 1 and model 2. System of equations with peso crisis defined as a dummy = 1 for 1994, 1995 and 1996

	Model 1	Model 2
= 1 if year \geq 1994 and year \leq 1996	0.031 (0.004)	0.047 (0.006)
Share skill 2	3.494 (0.738)	1.269 (0.239)
Share skill 3	1.975 (0.249)	0.465 (0.053)
Share skill 4	-0.044 (0.005)	-0.032 (0.003)
Share age 2	2.689 (0.569)	2.101 (0.366)
Share age 3	-0.322 (0.308)	0.012 (0.069)
Year	-0.027 (0.004)	-0.045 (0.006)
Constant	52.913 (8.347)	93.235 (12.479)
Observations	16	16

Standard errors in parenthesis. Sample of workers aged 25–60. Dependent variable: share of informal workers

Table 5 Regression of wage inequality on the share of informal workers in model 1 and in model 2. System of equations with peso crisis defined as a dummy = 1 for 1994 and 1995

	Model 1	Model 2
Informality	1.266 (0.348)	0.853 (0.367)
Share skill 2	1.199 (0.917)	-0.376 (0.360)
Share skill 3	0.993 (0.766)	0.395 (0.222)
Share skill 4	0.020 (0.016)	0.001 (0.013)
Share age 2	-1.720 (0.933)	-0.767 (0.791)
Share age 3	-1.613 (0.562)	-0.649 (0.150)
Year	-0.001 (0.007)	0.005 (0.014)
Constant	2.081 (13.565)	-10.837 (29.318)
Observations	16	16

Standard errors in parenthesis. Sample of workers aged 25–60. Dependent variable: Gini (model 1); var log wages (model 2)

Table 6 Regression of the share of informal workers on the peso crisis for model 1 and model 2. System of equations with peso crisis defined as a dummy = 1 for 1994 and 1995

	Model 1	Model 2
= 1 if year \geq 1994 and year \leq 1995	0.019 (0.006)	0.030 (0.009)
Share skill 2	1.752 (1.078)	0.764 (0.361)
Share skill 3	2.287 (0.433)	0.574 (0.087)
Share skill 4	-0.052 (0.009)	-0.038 (0.005)
Share age 2	1.676 (0.873)	1.625 (0.582)
Share age 3	0.552 (0.435)	0.195 (0.101)
Year	-0.020 (0.007)	-0.038 (0.009)
Constant	39.788 (13.156)	79.299 (20.195)
Observations	16	16

Standard errors in parenthesis. Sample of workers aged 25–60. Dependent variable: share of informal workers

Competing interests

The IZA Journal of Labor & Development is committed to the IZA Guiding Principles of Research Integrity. The author declares that she has observed these principles.

Acknowledgements

I thank Jackie Wahba and an anonymous referee for helpful comments and suggestions. All errors are mine. Responsible editor: Hartmut F. Lehmann

Received: 16 June 2015 Accepted: 12 January 2016

Published online: 10 March 2016

References

- Aleman-Castilla B (2006) The effect of trade liberalization on informality and wages: evidence from Mexico. *Cent Econ Perform Discus Pap No:763*
- Attanasio O, Goldberg PK, Pavcnik N (2004) Trade reforms and wage inequality in Colombia. *J Dev Econ* 74(2):331–366
- Bellù LG, Liberati P (2006) Policy impacts on inequality. Decomposition of income inequality by subgroups. *EASYPol Module:052*. http://www.fao.org/easypol/output/AbsLanguage.asp?id=444&job_no=00146&langa=EN
- Binelli C (2015) How the wage-education profile got more convex: evidence from Mexico. *B.E. J Macroecon* 12(2):509–560
- Binelli C, Attanasio O (2010) Mexico in the 1990s: the main cross-sectional facts. *Rev Econ Dyn* 13(1):1–14
- Binelli C, Rubio-Codina M (2013) The returns to private education: evidence from Mexico. *Econ Educ Rev* 36:198–215
- Bosch M, Maloney WF (2006) Gross worker flows in the presence of informal labor markets: the Mexican experience 1987–2002. *CEP Discussion Paper No:753*
- Bosch M, Manacorda M (2010) Minimum wages and earnings inequality in urban Mexico. *Am Econ J Appl Econ* 2:128–149
- Botero J, Djankov S, La Porta R, Lopez-de-Silanes F, Schleifer A (2004) The regulation of labor. *Q J Econ* 119(4):1339–1382
- Chong A, Gradstein M (2007) Inequality and informality. *J Public Econ* 91:159–179
- Davidson R, MacKinnon JG (1993) *Estimation and inference in econometrics*. Oxford University Press, New York
- Fitzsimons E, Malde B (2014) Empirically probing the quantity-quality model. *J Popul Econ* 27(1):33–68
- Galbraith JK (2012) *Inequality and instability*. Oxford University Press, New York
- Goldberg PK, Pavcnik N (2004) Trade, inequality, and poverty: what do I know? Evidence from recent trade liberalization episodes in developing countries. *Brookings Institute, Brookings Trade Forum 2004*
- Goldberg PK, Pavcnik N (2003) The response of the informal sector to trade liberalization. *J Dev Econ* 72(2):463–496
- Greene WH (2012) *Econometric analysis*. 7th ed. Prentice Hall, Upper Saddle River, NJ
- Hanson G, Harrison A (1999) Trade liberalization and wage inequality in Mexico. *Ind Labor Relat Rev* 52(2):271–288
- Johnson S, Kaufmann D, Zoido-Lobaton P (1998) Regulatory discretion and the unofficial economy. *Am Econ Rev Papers and Proc* 88(2):387–392
- Krstić G, Sanfey P (2010) Earnings inequality and the informal economy: evidence from Serbia. *Eur Bank Reconstr Dev Work Pap:114*
- Loayza NV (1996) The economics of the informal sector: a simple model and some empirical evidence from Latin America. *Cam-Roch Conf Ser Public Policy* 45:129–162

- Lopez-Acevedo G, Salinas A (2000) How Mexico's financial crisis affected income distribution. The World Bank, Washington D.C
- Maloney WF (2004) Informality revisited. *World Dev* 32(7):1159–1178
- Maloney WF (1999) Does informality imply segmentation in urban labor markets? Evidence from sectoral transitions in Mexico. *World Bank Econ Rev* 13(2):275–302
- Manacorda M, Sanchez-Paramo C, Schady NR (2010) Changes in returns to education in Latin America: the role of demand and supply of skills. *Ind Labor Relat Rev* 63(2):307–326
- Martin G (2000) Employment and unemployment in Mexico in the 1990s. *Monthly Labor Review* 123(11):3–18
- McKenzie D (2001) The household response to the Mexican peso crisis. Working Paper Department of Economics Stanford University. Available at: <http://www-siepr.stanford.edu/workp/swp01017.pdf>
- Meghir C, Narita R, Robin J-M (2015) Wages and Informality in Developing Countries. *Am Econ Rev* 105(4):1509–46
- Nevo A, Rosen A. (2012) Identification with imperfect instruments. *Rev Econ Stat* 94(3):659–671
- Pratap S, Quintin E (2006) The informal sector in developing countries. Output, assets and employment. UNU-WIDER Research Paper No. 2006 /130
- Robertson R (2004) Relative prices and wage inequality: evidence from Mexico. *J Int Econ* 64(2):387–409
- Rosser JB, Rosser MV, Ahmed E (2000) Income inequality and the informal economy in transition economies. *J Comp Econ* 28:156–171
- Rosser JB, Rosser MV, Ahmed E (2003) Multiple unofficial economy equilibria and income distribution dynamics in systemic transition. *J Post Keynesian Econ* 25(3):425–447
- Schneider F, Enste DH (2000) Shadow economies: size, causes, and consequences. *J Econ Lit* 38:77–114
- Shorrocks AF (1980) The class of additively decomposable inequality measures. *Econometrica* 48(3):613–625
- Verhoogen E (2008) Trade, quality upgrading and wage inequality in the Mexican manufacturing sector. *Q J Econ* 123(2):489–530
- Vuletin G (2009) What is the Size of the Pie? Measuring the Informal Economy in Latin America and the Caribbean. *Money Affairs* Vol. 21. pp 161–191
- Whitt JA (1996) The Mexican peso crisis. *Economic Review* Federal Reserve Bank of Atlanta, January/February

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Immediate publication on acceptance
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

Submit your next manuscript at ▶ springeropen.com
