

ECONOMIC  
RESEARCH  
F O R U M



مِنْتَدَى  
الْبَحْثِ  
الاقتصاديّة

2013

# working paper series

**HOW DIFFERENT ARE THE WAGE CURVES  
FOR FORMAL AND INFORMAL WORKERS?  
EVIDENCE FROM TURKEY**

**Badi H. Baltagi, Yusuf Soner Baskaya  
and Timur Hulagu**

**Working Paper No. 754**

# **HOW DIFFERENT ARE THE WAGE CURVES FOR FORMAL AND INFORMAL WORKERS? EVIDENCE FROM TURKEY**

Badi H. Baltagi, Yusuf Soner Baskaya and Timur Hulagu

**Working Paper 754**

**May 2013**

The views expressed in this paper are those of authors and do not necessarily reflect the official views of the Central Bank of Turkey. We thank Uwe Blien, Altan Aldan, Murat Kirdar, Zahid Samancioglu, Seref Saygili and Gonul Sengul as well as the participants at the 17th International Panel Data Conference held at McGill University in July 2011 for their comments and suggestions. We also benefited the comments and suggestions of three anonymous referees.

**Send correspondence to:**

Badi H. Baltagi

Department of Economics and Center for Policy Research, Syracuse University

[bbaltagi@maxwell.syr.edu](mailto:bbaltagi@maxwell.syr.edu)

First published in 2013 by  
The Economic Research Forum (ERF)  
21 Al-Sad Al-Aaly Street  
Dokki, Giza  
Egypt  
[www.erf.org.eg](http://www.erf.org.eg)

Copyright © The Economic Research Forum, 2013

All rights reserved. No part of this publication may be reproduced in any form or by any electronic or mechanical means, including information storage and retrieval systems, without permission in writing from the publisher.

The findings, interpretations and conclusions expressed in this publication are entirely those of the author(s) and should not be attributed to the Economic Research Forum, members of its Board of Trustees, or its donors.

## Abstract

This paper estimates wage curves for formal and informal workers using a rich individual level data for Turkey over the period 2005-2009. The wage curve is an empirical regularity describing a negative relationship between regional unemployment rates and individuals' real wages. While this relationship has been well documented for a number of countries including Turkey, less attention has focused on how this relationship differs for informal versus formal employment. This is of utmost importance for less developed countries where informal employment plays a significant role in the economy. Using the Turkish Household Labor Force Survey observed over 26 NUTS-2 regions, we find that real hourly wages of informal workers in Turkey are more sensitive to variations in regional unemployment rates than wages of formal workers. This is true for all workers as well as for different gender and age groups.

**JEL classification:** C26; J30; J60; O17

**Keywords:** Formal/Informal Employment; Wage Curve; Regional Labor Markets.

## ملخص

تقدر هذه الورقة منحنيات الأجور للعمال الرسمية وغير الرسمية باستخدام بيانات غنية على المستوى الفردي لتركيا خلال الفترة 2005-2009. منحني الأجور هو انتظام التجريبية التي تصف وجود علاقة سلبية بين معدلات البطالة الإقليمية والأجور الحقيقية للأفراد. وبينما توجد هذه العلاقة موثقة جيدا لعدد من الدول من بينها تركيا، إلا أن الانتباه كان بدرجة أقل لتحليل مدى اختلاف هذه العلاقة عن العمالة غير الرسمية مقابل العمالة الرسمية. هذا أمر في غاية الأهمية بالنسبة للبلدان الأقل نموا حيث تلعب العمالة غير الرسمية دورا هاما في الاقتصاد. باستخدام مسح القوى العاملة المنزلية التركية، نجد أن الأجور الحقيقية بالساعة للعمال غير الرسميين في تركيا هي أكثر حساسية للتغيرات في معدلات البطالة في المنطقة من أجور العمال الرسميين. وهذا صحيح بالنسبة لجميع العمال وكذلك لمختلف الجنسين والفئات العمرية.

## 1. Introduction

Informal employment is one of the key characteristics of developing country labor markets. Large numbers of workers in developing countries accept jobs with lower wages, worse working conditions and/or without access to social security coverage. According to a recent study by the OECD (2009), worldwide informal employment constitutes an average of 60% of total non-agricultural employment. This share varies from nearly 50% in Latin America, the Middle East and North Africa to almost 70% in South and Southeast Asia, and 75% in Sub-Saharan Africa. The majority of workers are expected to be still working in the informal sector in the foreseeable future, which highlights the importance of doing more research in this area, see Freeman (2010).

In this paper we estimate wage curves for formal and informal workers using a rich individual level dataset for Turkey over the period 2005-2009 from Turkstat Household Labor Force Survey. This survey provides information on a large set of demographic and job-related individual characteristics, including whether the individual is registered with the legally mandatory *Sosyal Güvenlik Kurumu* (Social Security Institution, SSI) at his current job.<sup>1</sup> Following Maloney (2004), OECD(2009) and Ramos et al. (2010), we define formal workers as those who are registered with the social security administration. During the sample period that we investigate, approximately 28 percent of wage earners in Turkey did not have the protection of the mandatory social security system.

Our empirical strategy is to estimate a wage curve for the formal and informal sectors in Turkey and to compare the unemployment elasticity of real wages across these groups. The empirical wage curve, pioneered by Blanchflower and Oswald (1990), estimates the sensitivity of individual wages to regional unemployment rates. Blanchflower and Oswald (1994, 2000) found that the unemployment elasticity of real wages in many countries is around -0.1, and dubbed it an empirical law, although a new generation of wage curve studies find different results.<sup>2</sup> For Turkey, Baltagi et al. (2012) find that the corresponding unemployment elasticity of real wages estimate is -0.099 over the period 2005-2008. This study uses the same data but updated for the period 2005-2009.<sup>3</sup>

Our study is also closely related to the literature comparing wage earnings for formal versus informal workers.<sup>4</sup> In fact, Baskaya and Hulagu (2011) show that for Turkey, informal workers earn approximately 15-20 percent less than their formal counterparts. In contrast to the literature focusing on the wage gap between formal and informal workers, this paper focuses on the sensitivity of formal and informal workers wages to regional unemployment variations in Turkey.<sup>5</sup>

Our findings indicate that workers without SSI coverage in Turkey have a larger unemployment elasticity of hourly wages. The magnitude of the difference between formal and informal workers varies across gender, education, experience and age groups. In almost all cases, we find that informal workers have a higher unemployment elasticity of hourly

---

<sup>1</sup>In Turkey, all workers must be registered in the Social Security Institution. This provides the workers with insurance for work-related accidents and illness, sickness, pregnancy, disability, old age and death as well as a retirement plan. Social security contributions are payable by workers and employers. Workers' contribution is 14 percent of gross salary, which is deductible in determining taxable income, whereas employee contribution is 19.5 percent.

<sup>2</sup>See Card (1995) for a critique as well as an excellent review of the first generation of wage curve studies.

<sup>3</sup>Although the Turkstat Household Labor Force Survey (THLFS) provides individual level data on a wide range of demographic and job-related characteristics starting from 2002, we focus on the post-2005 period, due to the absence of data on hours worked by the individuals for the 2002-2004 period.

<sup>4</sup>See Mazumdar (1976) for Peru; Mazumdar (1981) for Malaysia; Roberts (1989) and Gong and Van Soest (2001) for Mexico; Pradhan and Van Soest (1995) for Bolivia, Tansel (1999) for Turkey; and Pratap and Quintin (2006) for Argentina.

<sup>5</sup>The number of wage curves studies that focus on the formality status of the workers is relatively limited. See Berg and Contreras (2004) for an estimation of the informal wage curve for Chile, and Ramos et al. (2010) for Colombia.

wages. In addition, wages of informal workers who are younger, less educated or less experienced are more sensitive than those who are older, with more education or with more job experience. Moreover, our results reveal a huge variation across gender groups. In fact, we find that females in the informal labor markets face the highest sensitivity to regional unemployment variations.

Being one of the few studies in the literature focusing on potential differences in the wage curve relationships with respect to the formality status of workers, we think that this paper contributes to our knowledge on stylized characteristics of labor markets in developing countries with an emphasis on the observed wage flexibility. Nevertheless, it should also be acknowledged that these empirical findings may reflect various different mechanisms with different policy implications. For example, on the one hand, the unobserved productivity differences between formal and informal workers may lead to different degrees of wage flexibility. On the other hand, the informality status of the workers may itself be a factor resulting in higher wage flexibility. For example, if these workers without a formal protection face much lower wages during periods of a decline in the level of labor market activity. Identifying the exact mechanism behind these empirical findings is difficult and requires control for the selection issues due to unobserved individual heterogeneity. However, in the absence of a source of variation in our dataset which is irrelevant in the wage equation but would matter for the formality status of the individuals, this study is not able to disentangle the role of institutional factors from the role of unobserved individual heterogeneity across individuals.

The rest of the paper is structured as follows. Section 2 introduces the main features of our dataset. In section 3, we discuss our empirical strategy to estimate the unemployment elasticities of real hourly wages. Section 4 presents our empirical results concerning how informality of employment matters for the degree of wage flexibility both for the overall sample and with respect to worker types. Section 5 presents our concluding remarks.

## **2. Data**

The data used in this study is taken from the annual individual level data releases of the Turkstat Household Labor Force Survey (THLFS) for the 2005-2009 period.<sup>6</sup> As we are interested in how hourly wages of individuals respond to aggregate variations in the regional unemployment rates, we exclude unpaid family workers, self-employed individuals and individuals stated as employers in the survey. Also, due to possible measurement problems about their earnings, individuals younger than 15 years of age are excluded from the sample. Following the OECD (2009), we exclude the workers in the agricultural sector, as it is hard to distinguish between formal and informal employment in this sector. This yields 367,095 workers. In all regressions, we use the population weights provided by Turkstat.

Following the definition by OECD (2009), we categorize workers who are not registered by the SSI as informal workers.<sup>7</sup> The definition of informal employment by the International Labour Organization (ILO) includes not only the informal employees working in formal sector enterprises, but also all workers employed in informal sector enterprises and households producing goods exclusively for their own final use.<sup>8</sup> However, according to the McKinsey Global Institute (2003), most of the businesses in Turkey are registered and the informal economic activity mostly takes place in the form of partial reporting of revenues and employment. This provides another rationale for our choice for the informality measure.

---

<sup>6</sup>All private households who are living in the territory of the Republic of Turkey are covered by the THLFS. Residents of schools, dormitories, kindergartens, rest homes for elderly persons, special hospitals, military barracks and recreation quarters for officers are not covered. For more information, see the Turkstat website.

<sup>7</sup>This definition is also in line with various studies, such as Maloney (2004) and Ramos et al. (2010).

<sup>8</sup>See also Freeman (2010).

The data on hourly wages is obtained by dividing the monthly nominal after tax cash earnings, which exclude SSI contributions, by total hours worked in a month. We deflate the hourly wages into 2008 prices using regional price indices provided by Turkstat.<sup>9</sup> Finally, to measure the labor market activity, we use unemployment rates provided by Turkstat at the NUTS-2 level for 26 regions.<sup>10</sup>

### 3. The Model

Following the general practice in the literature, we estimate the wage curve relationship using a Mincerian wage regression extended to allow for the potential wage effects of variations in the aggregate regional unemployment rates.<sup>11</sup> As our main focus is on whether the wage curve relationship differs with respect to the formality status of the workers, we estimate this equation separately for formal and informal workers. Using the binary indicator  $f$ , equal to 1 for formal workers and 0 for informal workers, our statistical model becomes:

$$\log W_{irt} = \alpha^f + \beta^f \log U_{rt} + X'_{irt} \gamma^f + \mu_r^f + \lambda_t^f + v_{irt}^f, \quad f = 0,1 \quad (1)$$

where  $W_{irt}$  is the real hourly wage rate of individual  $i$  observed in region  $r$  at time  $t$ ,  $U_{rt}$  is the non-agricultural unemployment rate in region  $r$  at time  $t$ .  $X_{irt}$  represents the set of individuals' observed characteristics including age, gender, marital status, employment location, years of education, enrollment in school, years of tenure at the firm, firm size, industry of employment according to the NACE Rev. 1.1 classification, occupational group according to the ISCO-88 classification, permanency of the job and part-time work.<sup>12</sup> In order to account for the unobserved time-invariant regional differences and the shocks common across all regions within worker types, we also include region effects and year fixed effects denoted by  $\mu_r^f$  and  $\lambda_t^f$  respectively. Finally,  $v_{irt}^f$  is the remainder error term.

## 4. Empirical Results Regarding the Wage Curves for Formal and Informal Workers

### 4.1 Fixed Effects Results Treating Unemployment Rates as Predetermined

Table 1 presents the estimation results for the unemployment elasticity of real wages  $\beta$ , for all workers as well as formal and informal workers.<sup>13</sup> Panel A, denoted by FE, present the estimation results with region and time fixed effects, treating the regional unemployment rates as predetermined. The results show that the unemployment elasticity of real wages are low, yielding -0.043 for non-agricultural workers. Table 1 also finds that the elasticity for formal workers is lower than that for informal workers. This is -0.030 (formal) as compared to -0.059 (informal).

### 4.2 Accounting for Potential Endogeneity of Unemployment Rates

The preceding section does not consider the fact that the wages and unemployment rates are jointly determined. If the regional unemployment rate is endogenous, then the FE estimates yield biased and inconsistent estimates. Empirical evidence indicating the existence of such a bias in the FE estimator has been provided by Baltagi and Blien (1998) for Germany, Shilov

<sup>9</sup>Turkstat computes regional price indices by aggregating the elementary indices at the regional level, which are simple ratios of geometric mean of prices using the Laspeyres formula. In the aggregation process, the weights used by Turkstat for items in the consumption basket are proportional to expenditure shares obtained in 2004-2005-2006 Household Consumption Surveys.

<sup>10</sup>When calculating the NUTS-2 level regional unemployment rates, Turkstat includes both types of workers in terms of registry status to SSI.

<sup>11</sup>See Blanchflower and Oswald (1990, 1994, 2000) for pioneering work on the wage curve. See Baltagi et al. (2012) for recent evidence on the existence of a wage curve relationship in Turkey.

<sup>12</sup>These variables are explained in details in the Data Appendix.

<sup>13</sup>In order to save space, we only report  $\beta$ . However, the complete set of results are available upon request from the authors.

and Möller (2009) for Russia and more recently by Baltagi et al. (2012) for Turkey. Following Baltagi and Blien (1998), we use the one year lagged value of the unemployment rate as an instrument for the unemployment rate at time  $t$ . Figure 1 shows that there is a strong correlation, approximately 0.77, between the regional unemployment rates at time  $t$  and  $t - 1$ , suggesting that lagged unemployment rates by region is not a weak instrument for the contemporaneous value of the regional unemployment rates. We also formally reject the hypothesis that the lagged unemployment rate is a weak instrument for the contemporaneous unemployment rate using the Kleibergen-Paap F-statistic, which are significantly above the Stock-Yogo critical values (see Panel B of Table 1).

The estimation results in Panel B of Table 1 are in line with the earlier findings in the literature, as they suggest much higher elasticities (in absolute value) when the unemployment rates are not treated as predetermined. The results for this specification, denoted by FE-2SLS, indicate an elasticity of -0.107. This is consistent with what Blanchflower and Oswald (1994) regarded as an empirical regularity observed across different economies.

Panel B of Table 1 also finds that the estimated unemployment elasticity of hourly real wages for informal workers is more than 3 times larger than the estimated elasticity for formal workers. In fact, this elasticity is estimated at -0.071 for formal workers as compared to -0.263 for informal workers.

#### ***4.3 The Sensitivity of the Unemployment Elasticity of Real Wages for Formal and Informal Workers by Worker Types***

Table 2 presents the sensitivity of the unemployment elasticity of real wages for formal and informal workers by worker types. We find that the real hourly wages of females are more sensitive to variations in the unemployment rates than males (-0.156 as compared with -0.092). This is also true for younger workers as compared with older workers (-0.124 as compared with -0.059), and workers having less years of tenure compared with those having more years of tenure (-0.185 as compared with -0.034). On the other hand, the difference between workers with low and high levels of education is not as dramatic (-0.100 compared with -0.089).

Next, we analyze how informal and formal workers within different categories differ from each other with respect to the unemployment elasticity of real hourly wages. Panel A of Table 2 finds that there is a dramatic difference in the informal unemployment elasticity between males and females (-0.181 as compared to -0.505). Panel B of Table 2 finds that both formal and informal young workers are sensitive to variations in unemployment rates, with respective elasticities of -0.114 and -0.192. In contrast, among older workers, only the informal workers display a significant wage curve relationship. Panel C of Table 2 finds that wages of informal workers are more sensitive than their formal counterparts both in low and high tenure subgroups. Formal workers with high tenure had an insignificant wage curve. Finally, Panel D of Table 2 finds that the highest unemployment elasticity of real hourly wages is attained by informal workers with low education. Whatever worker type we considered, it is clear that the informal elasticity estimates are higher in absolute value than their formal counterparts.

#### ***4.4 The Wage Curves for Informal and Formal Workers with Similar Observable Characteristics***

The preceding sections find that the wage curves for informal workers are steeper than those for formal workers. One explanation for the steeper wage curves for informal workers is that they may also have less favorable observed traits such as lower years of schooling, no job tenure, different age, etc. (see Table A1), and less productivity, which is unobserved. The

THLFS survey lacks measures on individuals characteristics which would be irrelevant in the wage equation but yet matter for the informality status. It should be noted that we are not able to incorporate a correction in our Mincerian wage equation for selection into formality in a formal way, such as the one proposed by Heckman (1979). Therefore, in the absence for a formal answer for the degree to which the unobserved heterogeneity across informal and formal workers derive these results, we are not able to identify the exact mechanism leading to higher wage flexibility for the informal workers.

Having acknowledged that selection into the groups of informal and formal workers might bias our results, in this section, we use a two-step procedure for estimating wage curves for formal and informal workers who are similar in terms of their observed characteristics. In the first step, we estimate a probit model for predicting the probability of being a formal worker conditional on the individual's observed characteristics, such as age, education, gender, marital status, as well as job-related characteristics such as years of tenure, occupation, and industry. Using the predicted probabilities, the second step entails dividing the sample into two subsamples, i.e., individuals below and above the median predicted probability of being a formal worker. We then estimate the wage curves for formal and informal workers for these two subsamples separately.<sup>14</sup>

Table 3 presents the probit marginal effects estimates relating the formality status and the individual's observed characteristics, where the dependent variable takes the value of 1 for workers registered with the social security system and zero otherwise. The potential determinants of the formality status involves age, gender, job tenure, marital status, size of the firm where individuals are working, education level, occupation and industry of employment. These results are mostly in line with our expectations. To summarize, we find that the probability of being a formal worker increases monotonically with education level, the years of job tenure and age, where the difference in the propensities due to an extra year of job tenure and age decreases with the level of job tenure and ages. Males are more likely to be formal workers than females, conditional on other observed characteristics. We also find that married individuals are more likely to be formal workers.

The results presented in Table 4 show that informal and formal workers with predicted probability of being formal either below or above the median probability have different wage curves.<sup>15</sup> In particular, the unemployment elasticities of real wages for informal and formal workers in the subsample of predicted probabilities below the median are -0.267 and -0.177, respectively. Similarly, the elasticities for the informal and formal workers with a predicted probability of being a formal worker above the median is -0.317 and -0.037. The fact that informal workers have steeper wage curves than formal workers with similar predicted probabilities of being a formal worker suggest that the observed differences between formal and informal workers are not responsible for the differences in the slopes of the wage curves.

On the other hand, as noted above, these results should be tempered by the fact that any unobserved difference among individuals that would matter for the formality status may also

---

<sup>14</sup>In a sense, our estimation strategy can be regarded as using projected probabilities of being a formal worker conditional on observed traits in order to form subgroups of individuals with similar observed characteristics. Mechanically, the first stage of our exercise may resemble propensity score matching (PSM) since it entails predicting probability of being a formal worker conditional on observed characteristics. However, it should be noted that it differs from PSM which would entail measuring the effect of treatment effect of a policy, program or a status by comparing the outcomes of individuals with similar probabilities of receiving a treatment (conditional on observed characteristics) but different treatment status. In our setting, our exercise could be regarded as PSM if we were to compare the sensitivity of individuals' wages to local unemployment rates, i.e.  $\beta_i^f$  values, with different formality status but similar probability of being formal using semi-parametric matching techniques. However, since such  $\beta_i^f$  values are not available, this exercise is not feasible.

<sup>15</sup>We also reach similar results when we define the subsamples with respect to deciles, quintiles or quartiles of the distribution of the predicted formality probabilities.

affect the sensitivity of their wages to the unemployment variations. Therefore, the evidence provided in this paper, limited with what is available in our dataset, is not sufficient to identify the exact mechanisms leading to higher sensitivity of wages of informal workers to aggregate labor market fluctuations.<sup>16</sup>

## 5. Conclusion

Using a rich individual level data set for Turkey observed over the period 2005-2009, this paper estimated wage curve relationships with a particular focus on whether informal and formal workers differ from each other with respect to the sensitivity of real hourly wages to unemployment rates. We find that the unemployment elasticity of real hourly wages for informal workers is around -0.25. It is remarkable that this figure is 2.5 times the estimated elasticity obtained for all workers and 3.5 times the estimate for formal workers. This is even more dramatic by worker type, such as male versus female. No matter what worker type is used, the informal elasticity estimates are higher in absolute value than their formal counterparts. We also show that steeper wage curves for informal workers hold even within the sample of individuals who have similar observable characteristics as implied by the predicted probabilities of being a formal worker.

We believe that these higher elasticities and the stronger wage curve relationship for the informal workers contribute to our knowledge on the stylized characteristics of the labor markets in developing countries, as the number of studies focusing on how the wage curve relationship differs for informal and formal workers are quite limited despite the substantial share of informal employment in such countries. However, failing to account for the selection into the groups of informal and formal workers, we acknowledge the need for further research which should focus on identifying the role of different factors, such as unobserved productivity differences or the institutional differences between formal and informal labor markets with respect to job security and/or wage determination mechanisms, that would lead to steeper wage curves for the informal workers.

---

<sup>16</sup>We thank one of the referees for raising this important issue.

## References

- Baltagi, BH, YS Baskaya and T. Hulagu. 2012. The Turkish wage curve: Evidence from the household labor force survey. *Economics Letters* 114: 128-131
- Baltagi, BH and U. Blien. 1998. The German wage curve: evidence from the IAB employment sample. *Economics Letters* 61: 135-142
- Baskaya, YS, and T. Hulagu. 2011. Informal-Formal worker wage gap in Turkey: Evidence from a semi-parametric approach. *Central Bank of Turkey Working Paper* 11/15
- Berg, J. and D. Contreras. 2004. Political-Economic regime and the wage curve: Evidence from Chile, 1957-96. *International Review of Applied Economics* 18(2): 151-165
- Blanchflower, DG and AJ. Oswald. 1990. The wage curve. *Scandinavian Journal of Economics* 92: 215-235
- \_\_\_\_\_. 1994. *The Wage Curve*. MIT Press, Cambridge MA
- \_\_\_\_\_. 2000. International wage curves. In: Freeman R and Katz L (eds) *Differences and Changes in Wage Structures*. University of Chicago Press and NBER
- Card, D. 1995. The wage curve: a review. *Journal of Economic Literature* 33: 285-299
- Freeman, RB. 2010. Labor regulations, unions, and social protection in developing countries: Market distortions or efficient institutions? In: Rodrik D and Rosenzweig M (eds) *Handbook of Development Economics* 5: 4657-4702. Elsevier B.V. Amsterdam, North Holland
- Gong, X and A. Van Soest. 2001. Wage differentials and mobility in the urban labor market: A panel data analysis for Mexico. *Labour Economics* 9(4): 513-529
- Heckman, J. 1979. Sample selection bias as a specification error. *Econometrica* 47: 153-161
- Kleibergen, F. and R. Paap. 2006. Generalized reduced rank tests using the singular value decomposition. *Journal of Econometrics* 133: 97-126
- Maloney, WF. 2004. Informality revisited. *World Development* 32: 1159-1178
- Mazumdar, D. 1976. The urban informal sector. *World Development* 4: 655-679
- \_\_\_\_\_. 1981. *The urban labor market income distribution: A study of Malaysia*. Oxford University Press, Oxford
- Mincer, J. 1974. *Schooling, experience and earnings*. Columbia University Press, New York
- McKinsey Global Institute. 2003. *Turkey: Making the Productivity and Growth Breakthrough*. McKinsey, New York
- OECD. 2009. Is informal normal? Towards more and better jobs in developing countries. In: Jutting J and de Laiglesia JR (eds) *OECD Development Centre Studies*, Paris.
- Pradhan, M. and A. Van Soest. 1995. Formal and informal sector employment in urban areas of Bolivia. *Labour Economics* 2: 275-297
- Pratap, S. and E. Quintin. 2006. Are labor markets segmented in developing countries? A semiparametric approach. *European Economic Review* 50: 1817-1841
- Ramos, R., JC. Duque, and J. Surinach. 2010. Is the wage curve formal or informal? Evidence for Colombia. *Economics Letters* 109: 63-65
- Roberts, BR. 1989. Employment structure life cycle and life chances: Formal and informal sectors in Guadalajara. In: Portes A, Castells M, Benton LA (eds) *The Informal*

*Economy: Studies in Advanced and Less Developed Countries*. Johns Hopkins University Press, Baltimore.

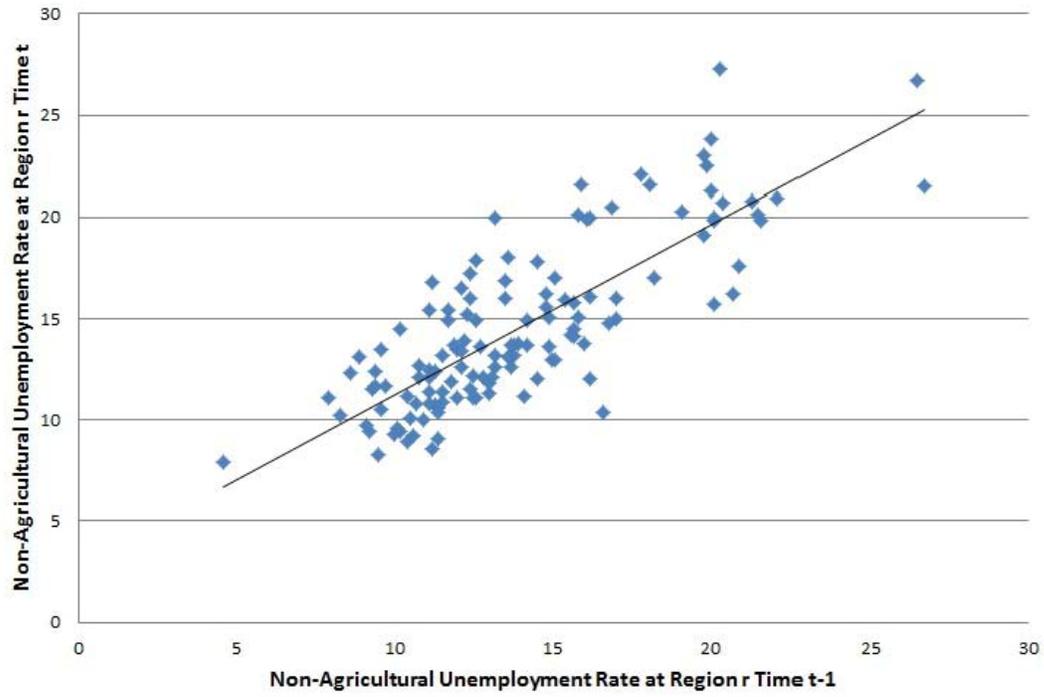
Shilov, A. and J. Möller. 2009. The wage curve in Russia, 1995-2005. *Economics Letters* 102: 90-92

Stock, JH and M. Yogo. 2005. Testing for weak instruments in linear IV regression. In: *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, 80-108. Andrews DWK and Stock JH (eds) Cambridge University Press, Cambridge

Tansel, A. 1999. Formal versus informal sector choice of wage earners and their wages in Turkey. *Economic Research Forum Working Paper* No. 9927

Türkiye İstatistik Kurumu (Turkish Statistical Institute, Turkstat). 2009. *Household Labor Force Survey Data 2005-2008*

**Figure 1: Persistency in Regional Unemployment Rates**



Notes: Data points reflect unemployment rates by region and time, which are available from Turkstat.

**Table 1: The Unemployment Elasticity of Real Hourly Wages By Formality/Informality Status**

	<b>All</b>	<b>Formal</b>	<b>Informal</b>
Panel A: FE	-0.043 (0.008)***	-0.030 (0.009)***	0.059 (0.018)***
$R^2$	0.598	0.584	0.304
	<b>All</b>	<b>Formal</b>	<b>Informal</b>
Panel B: FE-2SLS	-0.107 (0.020)***	-0.071 0.021***	-0.263 (0.045)***
$R^2$	0.598	0.584	0.303
K-P F stat.	5.0e04	3.7e04	1.1e04
Obs.	367,095	274,495	92,600

Notes: a) Robust standard errors in parentheses. \*, \*\* and \*\*\* represent significance at 10%, 5% and 1%, respectively. b) In FE-2SLS specification, the logarithm of non-agricultural unemployment rate by region in the previous year is used as an instrument for the logarithm of non-agricultural unemployment rate by region at time t. c) The set of observed characteristics which are used to control for individual heterogeneity include age, gender, marital status, employment location, years of education, enrollment to a school, years of tenure at the firm, firm size, industry of employment according to the NACE Rev. 1.1 classification, occupational group according to the ISCO-88 classification, permanency of the job and part-time work. d) Robust Kleibergen-Paap Wald rk F statistics suggest that the hypothesis that lagged value of regional unemployment rates is a weak instrument for contemporaneous value of regional unemployment rates is rejected, considering the Stock-Yogo "rule-of-thumb" critical value for weak instrument is 10.

**Table 2: The Unemployment Elasticity of Real Hourly Wages by Formality/Informality Status and Worker Types**

PANEL A: Gender						
	All	Male Formal	Informal	All	Female Formal	Informal
FE-2SLS	-0.092 (0.022)***	-0.061 (0.024)**	-0.181 (0.046)***	-0.156 (0.042)***	-0.113 (0.039)***	-0.505 (0.143)***
R <sup>2</sup>	0.585	0.569	0.324	0.653	0.642	0.344
K-P F stat.	5.2e04	3.9e04	1.2e04	1.6e04	1.3e04	0.2e04
Obs.	286,034	212,916	73,118	81,061	61,579	19,482

PANEL B: Age						
	All	Old Formal	Informal	All	Young Formal	Informal
FE-2SLS	-0.059 (0.028)**	-0.022 (0.029)	-0.285 (0.075)***	-0.124 (0.027)***	-0.114 (0.029)***	-0.192 (0.056)***
R <sup>2</sup>	0.591	0.583	0.281	0.560	0.538	0.265
K-P F stat.	3.2e04	2.6e04	0.6e04	3.5e04	2.6e04	0.9e04
Obs.	167,645	130,656	36,989	199,450	143,839	55,611

PANEL C: Tenure						
	All	High Formal	Informal	All	Low Formal	Informal
FE-2SLS	-0.034 (0.027)	-0.005 (0.028)	-0.266 (0.082)***	-0.185 (0.027)***	-0.141 (0.030)***	-0.298 (0.053)***
R <sup>2</sup>	0.595	0.551	0.285	0.509	0.497	0.305
K-P F stat.	3.0e04	2.5e04	0.4e04	3.8e04	2.7e04	1.1e04
Obs.	137,393	119,617	17,776	229,702	154,878	74,824

PANEL D: Education						
	All	High Formal	Informal	All	Low Formal	Informal
FE-2SLS	-0.089 (0.026)***	-0.073 (0.026)***	-0.136 (0.094)	-0.100 (0.030)***	-0.038 (0.034)	-0.288 (0.051)***
R <sup>2</sup>	0.583	0.538	0.426	0.374	0.347	0.256
K-P F stat.	3.7e04	3.3e04	0.3e04	3.0e04	1.9e04	1.1e04
Obs.	178,755	158,767	19,988	188,340	115,728	72,612

Notes: a) Robust standard errors in parentheses. \*, \*\* and \*\*\* represent significance at 10%, 5% and 1%, respectively. b) Young (old) refers to individuals younger (older) than sample mean value for years of age, which is 34.1. Low (high) tenure refers to individuals with tenure less (more) than the sample mean value, which is 6.94 years. Low (high) education refers to individuals with less than or equal to 8 years of schooling (more than 8 years of schooling). c) In FE-2SLS specification, the logarithm of non-agricultural unemployment rate by region in the previous year is used as an instrument for the logarithm of non-agricultural unemployment rate by region at time t. d) The set of observed characteristics which are used to control for individual heterogeneity include age, gender, marital status, employment location, years of education, enrollment to a school, years of tenure at the firm, firm size, industry of employment according to the NACE Rev. 1.1 classification, occupational group according to the ISCO-88 classification, permanency of the job and part-time work. e) Robust Kleibergen-Paap Wald rk F statistics suggest that the hypothesis that lagged value of regional unemployment rates is a weak instrument for contemporaneous value of regional unemployment rates is rejected, considering the Stock-Yogo "rule-of-thumb" critical value for weak instrument is 10. tablenotes

**Table 3: Determinants of Formality Status**

Age	0.035 (0.001)***
Age <sup>2</sup>	-0.0005 (0.000)***
Female	-0.058 (0.002)***
Single	0.024 (0.005)***
Married	0.069 (0.005)***
Urban	-0.004 (0.002)***
Enrolled	0.048 (0.003)***
Permanent	0.302 (0.005)***
Part-time	-0.219 (0.008)***
Experience	0.019 (0.000)***
Experience <sup>2</sup>	-0.0004 (0.000)***
Firm Size (10-24 Emp)	0.108 (0.001)***
Firm Size (25-49 Emp)	0.146 (0.001)***
Firm Size (50-249 Emp)	0.200 (0.001)***
Firm Size (250-499 Emp)	0.147 (0.001)***
Firm Size (Emp > 500)	0.161 (0.001)***
Primary school grads	0.087 (0.003)***
Secondary school grads	0.099 (0.002)***
High school grads	0.157 (0.003)***
University grads	0.174 (0.002)***
Number of Observations	367,095

Notes: a) Marginal effects; Estimated with probit to predict the propensity scores. b) Other control variables include occupation and industry dummies. c) Omitted categories for marital status, firm size and education are divorced or widowed, less than 10, and not completed any educational institution, respectively. d) The numbers in parentheses are robust standard errors clustered for within region correlations and \*, \*\* and \*\*\* represent significance at 10%, 5% and 1%, respectively.

**Table 4: The Unemployment Elasticity of Real Hourly Wages By Formality/Informality Status and Propensity Scores**

	<b>Formal</b>	<b>Informal</b>
Above median	-0.037 (0.024)	-0.317 (0.199)
Below median	-0.177 (0.038)***	-0.267 (0.046)***

Notes: a) Results of FE-2SLS specification are given. The logarithm of non-agricultural unemployment rate by region in the previous year is used as an instrument for the logarithm of non-agricultural unemployment rate by region at time t. b) Robust standard errors in parentheses. \*, \*\* and \*\*\* represent significance at 10%, 5% and 1%, respectively. c) The set of observed characteristics which are used to control for individual heterogeneity include age, gender, marital status, employment location, years of education, enrollment to a school, years of tenure at the firm, firm size, industry of employment according to the NACE Rev. 1.1 classification, occupational group according to the ISCO-88 classification, permanency of the job and part-time work.

## Appendix

### 6. Data Appendix

In this appendix, we provide details about our dataset. First, we summarize our data with respect to informality for different subgroups. Particularly, Table 5 lists percentages of formal and informal workers for four individual characteristic categorizations and three different samples. In Table 6, we present means and standard deviations of wage levels, age, education levels, experience, the rate of permanent and part-time jobs for both formal and informal workers.

We also give details about the individual specific control variables that we use. These are listed below:

- Age. The survey provides eleven age categories in 5-year intervals.
- Gender. Female=1 and Male=0.
- Marital status. Married=1, and zero otherwise.
- Employment location. Urban=1 and Rural=0.
- Education. The variable *Education* is years of completed education, while the variable *Enrolled* is a binary variable which takes the value 1 for individuals enrolled to a school, and zero otherwise.
- Social security registration: Binary variable which takes the value 1 if the individual is registered in the social security administration, and zero otherwise.
- The individual's years of tenure at the firm. This is calculated as the starting year at the current job subtracted from the survey year.
- Industry classification. This is a set of seven binary variables categorized according to the NACE Rev. 1.1 classification pertaining to the non-agricultural industries, which include mining, manufacturing, electricity, construction, transportation, trade, finance, and community, social and personal services.
- Occupational group. This is a set of seven binary variables defined in line with non-agricultural occupations categorization according to the ISCO-88 classification, which include legislators, senior officials and managers; professionals; technicians and associate professionals; fishery workers; clerks; service workers and shop and market sales workers; craft and related trades workers; plant and machine operators and assemblers; and elementary occupations.
- Permanency of the job. This is a set of three mutually exclusive binary variables describing whether the job is *permanent*, *temporary* or *seasonal*.
- Employment type. Full-time=0 and part-time=1.
- Other activity to earn income. Yes=1 and no=0.
- Firm size. This is measured by the number of persons employed in the firm and summarized by six binary variables corresponding to the following categories: less than 10 employees, 10-24, 25-49, 50-249, 250-499, and 500 and more.
- Employment status in the same month of last year. Binary variable which takes the value 1 if the individual was working in the same month of last year, and zero otherwise.

**Table 5: The Formality Percentages of Workers by Types**

	<b>Formal (%)</b>	<b>Informal (%)</b>
Gender		
Male	74.44	25.56
Female	75.97	24.03
Age		
Old	77.94	22.06
Young	72.12	27.88
Tenure		
High	87.06	12.94
Low	67.43	32.57
Education		
High	88.82	11.18
Low	61.45	38.55

Source: Authors' Calculations Using THLFS. Workers in the agricultural sector are excluded. Young (Old) refers to individuals younger (older) than sample mean value for years of age, which is 34.1. Low (high) tenure refers to individuals with tenure less (more) than the sample mean value, which is 6.94 years. Low (high) education refers to individuals with less than or equal to 8 years of schooling (more than 8 years of schooling).

**Table 6: Summary Statistics for Some Characteristics by Formality Status**

	<b>Formal</b>	<b>Informal</b>	<b>Total</b>
Real wage	5.14 (5.45)	2.31 (3.19)	4.37 (5.10)
Age	34.67 (9.32)	32.82 (12.64)	34.16 (10.36)
Education	9.91 (3.89)	6.75 (2.95)	9.05 (3.92)
Experience	7.84 (7.65)	4.24 (6.80)	6.86 (7.60)
Permanency	97.41%	68.31%	89.49%
Part-time rate	1.50%	5.91%	2.70%

Source: Authors' Calculations Using THLFS. Workers in the agricultural sector are excluded. Means of characteristics are reported while standard errors are given in parentheses.