DISCUSSION PAPER SERIES

IZA DP No. 9455

Firm-Size Wage Gaps along the Formal-Informal Divide: Theory and Evidence

Binnur Balkan
Semih Tumen

October 2015

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Firm-Size Wage Gaps along the Formal-Informal Divide: Theory and Evidence

Binnur Balkan
Central Bank of the Republic of Turkey

Semih Tumen
Central Bank of the Republic of Turkey and IZA

Discussion Paper No. 9455
October 2015

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

IZA Discussion Paper No. 9455
October 2015

ABSTRACT

Firm-Size Wage Gaps along the Formal-Informal Divide: Theory and Evidence*

Observationally equivalent workers are paid higher wages in larger firms. This fact is often named as the "firm-size wage gap" and is regarded as a key empirical puzzle. Using micro-level data from Turkey, we document a new stylized fact: the firm-size wage gap is more pronounced for informal (unregistered) jobs than for formal (registered) jobs. To explain this fact, we develop a two-stage wage-posting game with market imperfections and segmented markets, the solution to which produces wages as a function of firm size in a well-defined subgame-perfect equilibrium. The model proposes two explanations. *First*, taxes on formal employment generate a wedge between formal and informal size wage gaps. Thus, government policy can potentially affect the magnitude of the firm-size wage gaps. The *second* explanation features a market-based framework with strategic interactions. Relative to small firms, large firms typically post higher wages for both formal and informal jobs they open. A high-wage formal job attracts a larger pool of applicants than a high-wage informal job. The larger pool of applicants for the formal job, in turn, allows the firm to somewhat lower the initial wage offer, while this second-round effect is negligible for informal jobs. As a result, size differentials are lower in formal jobs than informal jobs. We argue that the observed patterns in the use of social connections in job search and heterogeneity in job preferences can be used to justify the validity of this second mechanism.

JEL Classification: C78, J21, J31, L11

Keywords: firm size, wage gap, informal job, wage posting, subgame perfection, taxes, social networks

Corresponding author:

Semih Tumen
Research and Monetary Policy Department
Central Bank of the Republic of Turkey
Istiklal Cad. No:10
06100 Ulus, Ankara
Turkey
E-mail: semih.tumen@tcmb.gov.tr

* We thank Daron Acemoglu, Matteo Bobba, Kerem Cosar, Hakan Ercan, Ercan Uygur, seminar participants at the Central Bank of the Republic of Turkey, the participants of the Turkish Economic Association Workshop on Economic Statistics in Ankara, BETAM/World Bank Labor Market Network Meeting in Istanbul, International Economic Association Conference in Dead Sea/Jordan, IZA/World Bank Conference on Employment and Development in Lima, and Turkish Economic Association Conference in Antalya for useful suggestions. We are particularly grateful to Kitt Carpenter (the editor) and two anonymous referees for very helpful comments. The views expressed here are of our own and do not necessarily reflect those of the Central Bank of the Republic of Turkey. All errors are ours.

1 Introduction

It is well-documented in the literature that larger firms pay higher wages to observationally equivalent workers than smaller firms pay.[1] This fact holds almost invariably across countries and sectors as well as across jobs with different supervisory responsibilities.[2] Several explanations are offered in the literature ranging from unobserved worker heterogeneity [Idson and Feaster (1990)] to unobserved firm productivity [Idson and Oi (1999)] and from the need for better data [Troske (1999)] to firm-level differences in labor turnover due to differences in hiring and human resource management practices [Idson (1996)]. Still, the firm-size wage gap is regarded as one of the key empirical challenges in labor economics and additional research is called for to enhance our understanding of this observed phenomenon.

In this paper, we document a new fact: the firm-size wage gap is higher for informal jobs than formal jobs. We perform our empirical analysis using a nationally-representative micro-level dataset from Turkey, which we believe that is a good laboratory to study this question—since around 25 percent of all jobs are informal based on official figures. "Size" corresponds to the number of workers employed in a particular firm. The data allows us to define firm size in 6 categories, 1 being the smallest and 6 being the largest. After controlling for a comprehensive set of observed covariates, we find that the wage gap between the firms of the largest versus the smallest size (i.e., Size 6 versus Size 1) is 33.2 percent and 46.7 percent for formal and informal jobs, respectively. This means that the average pay differential between the largest and the smallest firm is approximately 14 percent higher in informal jobs than that for formal jobs. Moreover, this difference increases monotonically—from 0 to 14 percent—as size increases from 1 to 6; that is, when we consider wages as a function of size, our finding means that the slope of this function is steeper for informal jobs than formal jobs. See Figure (2.1) for a clear empirical visualization of this statement.

[1] See Oi and Idson (1999) for a comprehensive review of the early literature. Breakthrough papers in the early literature that deserve attention include Mellow (1982), Brown and Medoff (1989), and Groshen (1991).

[2] For studies documenting firm-size wage gaps at the country level, see, for example, Marcouiller, Ruiz de Castilla, and Woodruff (1997) for El Salvador, Mexico, and Peru, Tan and Batra (1997) for Colombia, Mexico, and Taiwan (China), Brunello and Colussi (1998) for Italy, Hollister (2004) for the United States, and Lallemand, Plasman, and Rycx (2007) for Belgium, Denmark, Ireland, Italy, and Spain. Baker, Jensen, and Murphy (1988) document sectoral differences in size-wage gaps for CEOs. Meagher and Wilson (2004) and Fox (2009) find that the size-wage gap is larger for jobs with managerial responsibilities.

Then comes the question: is it possible to develop a coherent theoretical framework to understand the forces driving this result? We construct a two-stage wage-posting game with market imperfections and segmented markets. The solution of this game analytically characterizes wages as a function of firm size within a well-defined subgame-perfect equilibrium. Firms differ in size, but workers are homogeneous. Each firm posts a wage, workers observe all offers and device a symmetric application strategy. Large firms offer higher wages, because the vacancies posted by them are more valuable—as larger firms are more productive. This framework proposes two mechanisms as potential explanations for the new stylized fact we report. *First*, formal jobs are subject to taxes and informal jobs are not. In our theoretical setup, taxes impose a wedge between the size premium in informal jobs versus that in formal jobs; that is, the size-wage gap is lower for formal jobs than informal jobs, because formal jobs are more costly to the employer due to taxes. We present auxiliary empirical evidence supporting this theoretical prediction. We conclude that government policy may affect the magnitude of the firm-size wage premium.

The model is able to propose a *second* explanation even when the taxes are shut down. A large firm faces a key tradeoff. It has to post a high enough wage so as to guarantee that the productive position is filled. There is also a secondary mitigating effect. The firm also has to keep wages at reasonable levels, because higher wages will attract a lot of applicants, the extent of which will provide incentives to keep the wage offer somewhat lower (as the vacancy will be filled anyway with that many applicants). We argue that the high-wage informal jobs posted by larger firms does not receive that many applications; so, the secondary (mitigating) force is weak for those jobs. As a result, the size gradient of the wage function is steeper for informal jobs than formal jobs. The key point is that the mitigating force is weaker for informal jobs. Why is this the case? One explanation is related to the use of social networks in job search. By definition, information on informal job openings most likely spread through social contacts (i.e., friends, relatives, and acquaintances). This fact, by itself, can generate the theoretical prediction that the applicant pool for a high-wage informal job is likely smaller than that of a formal job with comparable pay. An alternative explanation is that formal jobs

might be more desirable, since they offer better non-pecuniary benefits. If the non-pecuniary aspects of a job are valued by applicants, then it is understandable that a formal job might receive a larger pool of applicants than an informal job with similar pay. Finally, informal jobs are risky (i.e., they are terminated if caught by the authorities); as a result, risk averse applicants might try to avoid interruptions in job spells.

The second mechanism is also consistent with an equally intuitive explanation, which is the flip side of the one proposed above. It can also be argued that the wage penalty for working in a small firm is bigger for informal jobs than formal jobs, because the pool of applicants is larger for informal than formal low-wage jobs. Given that informal workers are more fragile (i.e., immigrant/low-educated/non-unionized), their bargaining power tends to be lower than that for formal workers. This suggests that the high negative wage penalty associated with informal jobs in small firms might be a consequence of a monopsonistic labor market, i.e., an upward sloping labor supply curve to an individual employer that would be steeper for informal than formal jobs [see, e.g., Manning (2003), Fakhfakh and Fitzroy (2006), and Mortensen (2010)].

Our theoretical model is similar to the wage-posting models of Montgomery (1991), Lang (1991), and Lang, Manove, and Dickens (2005). In line with these papers, we solve a two-stage wage-posting game in a subgame-perfect equilibrium. Different from them, we incorporate two pieces: (*i*) firms differ in size and (*ii*) the number of applicants for each position is an indirect function of firm size. Other than these two major differences, we adopt the idea that formal and informal jobs are posted in segmented markets. In this respect, our model is related to the dual labor markets literature arguing that informal and formal jobs are subject to market segmentation at least partially [see, for example, Stiglitz (1976), Dickens and Lang (1985), and Heckman and Hotz (1986)].[3]

There are several other literatures that our paper is directly related to. The first one is the literature offering theoretical explanations to firm-size wage gaps. Manning (2003) argues that at least part of the firm-size wage premia should be associated with the upward sloping

[3]See Magnac (1991) for an opposing view.

labor supply curve to an individual firm. In other words, he suggests that the monopsonistic structure in labor markets might be one plausible explanation. Lazear (2009) shows that turnover in larger firms are lower, since these firms facilitate the accumulation of firm-specific human capital more than smaller firms. As a result, firm-size wage gaps might be partly related to steeper—and uninterrupted—wage growth paths in large firms. According to this explanation, firm-size wage gap is a result of differential human capital investment patterns in large versus small firms. Garicano and Rossi-Hansberg (2006) construct a hierarchical matching model and find that larger firms attract more able managers, these managers amplify the productivity of the labor input, and, thus, firm-size premia emerge. Their explanation relates firm-size wage gaps to differences in organizational and human resource management practices across small and large firms. Finally, there are several papers suggesting that large firms are more likely to operate in imperfectly competitive markets, i.e., they have larger market power than small firms [see, e.g., Albaek, Arai, Asplund, Barth, and Strojer Madsen (1998)]. As a result, large firms might be able to accumulate more monopoly rents that they can share with their workers through bargaining than small firms [Mellow (1982)]. So, market power in the product market along with wage determination via bargaining/unions might also provide a viable mechanism.

The second literature is the one documenting significant wage differentials with respect to the formality status of the workers. For example, Mazumdar (1981), Heckman and Hotz (1986), Pradhan and van Soest (1995), Tansel (1999), and Gong and van Soest (2002) document that significant formal-informal wage gaps exist in Malaysian, Panamian, Bolivian, Turkish, and Mexican labor markets, respectively. Although some opposing views exist [see, e.g., Pratap and Quintin (2006)], formal-informal wage gaps are widely-agreed features of labor markets in developing countries. We would like to note that, in this paper, we are dealing with the formal-informal differences in the size gradient of the wage function, rather than differences in levels of average wages in these two sectors—although our analysis also captures the latter fact.

This paper makes several contributions to the literature on the firm-size wage gaps. To start

with, this is the first paper in the literature documenting that the firm-size wage gaps differ across informal and formal jobs. Around 25 percent of the working population in Turkey are employed in informal jobs; thus, the Turkish data offers a natural environment to investigate the differences in firm-size wage gap patterns between formal and informal jobs. Second, we develop a theoretical model to explain this phenomenon. One of the predictions of the model is that the firm-size wage gap is negatively related to the tax burden of formal jobs. In other words, the employers' tax burden is a potential determinant of the magnitude of the firm-size wage gaps. We confirm the validity of this prediction using state-level labor tax differences in the US. Finally, we argue—as the second prediction of our model—that the number of new applicants that an incremental increase in the wage offer will attract is smaller in informal jobs. This, itself, can explain the fact we document.

The plan of the paper is as follows. Section 2 provides summary statistics for our micro-level data from Turkey, describes the institutional environment, and presents the results of our empirical investigation. Section 3 constructs the benchmark model, solves it, assesses its main predictions, discusses potential policy implications, and performs additional empirical tests of the model with auxiliary data. Section 4 concludes.

2 Empirical Analysis

2.1 Data Description and Summary Statistics

We use the Turkish Household Labor Force Survey (THLFS) data collected and compiled by the Turkish Statistical Institute (TURKSTAT). This is a large, micro-level, survey-based, publicly-available, and nationally-representative dataset based on which the official unemployment and earnings figures have been calculated and published regularly. The micro-level data are publicly available only with yearly frequency, so we use yearly data from 2006 to 2012. The sample we focus on consists of employed individuals of age 15 and above. The wage variable describes after-tax hourly wage earnings in the main job and it is deflated using the GDP

	Informal		Formal	
Variable	**Mean**	**Std.Dev**	**Mean**	**Std.Dev**
Size 1	0.712	0.453	0.212	0.409
Size 2	0.118	0.322	0.126	0.331
Size 3	0.098	0.298	0.212	0.409
Size 4	0.058	0.234	0.274	0.446
Size 5	0.009	0.092	0.072	0.259
Size 6	0.006	0.074	0.104	0.306
Female	0.235	0.424	0.233	0.423
Married	0.581	0.493	0.711	0.453
Full-time	0.933	0.696	0.983	0.731
No Degree	0.121	0.326	0.015	0.123
Primary School	0.423	0.494	0.261	0.439
Secondary School	0.257	0.437	0.143	0.350
High School	0.091	0.287	0.141	0.348
Vocational High School	0.071	0.256	0.151	0.358
College & Above	0.038	0.191	0.289	0.453
Age	32.63	12.599	34.31	9.101
Log Real Wages (hourly)	0.971	0.614	1.708	0.632
Sample share	0.241		0.759	
# of Observations	139,888		441,866	

Table 1: **Summary Statistics:** Turkish Household Labor Force Survey data between 2006–2012 are used in the analysis. Wages are deflated taking 2012 as the base year. The first two columns describe the summary statistics for informal employment and the last two columns describe those for formal employment. Total number of observations is 581,754. Appropriate frequency weights are used.

deflator with 2012 being the base year.[4] Wages and hours of work are reported by the survey respondents—rafter than other sources, such as the employers and administrative records. Wages are reported as monthly wages in the survey, while hours worked are reported as hours worked per week. To convert the monthly wages into hourly wages, we follow the conventions in the literature and employ the formula (monthly wages/(weekly hours $\times$ 4.3)). The dataset features 27 occupation and 87 industry categories classified based on the standardized NACE Rev.2 rules. Moreover, we include 26 regional dummy variables—at the NUTS2 level—to capture potential regional variations in the firm-size wage gap patterns.

The THLFS questionnaire asks the size of the firm that each worker is employed. So, the unit of analysis for the size variable is "firm"—rather than establishment.[5] Our key variable, firm size, is defined by 6 dummy variables. We name these variables from Size 1 through Size

[4]There is no information on before-tax wage earnings in the dataset. The wage earnings does not include extra hours, performance bonuses, and other premia. Although, the THLFS asks about these side payments, there are many missing cells, which makes the use of this variable infeasible. So, we focus only on after-tax hourly earnings excluding bonuses and other pecuniary benefits.

[5]See Barth and Dale-Olsen (2011) for a description of the conceptual differences between firm- and establishment-size wage gaps.

6, the latter being the largest firm. The Size 1 firm is a firm with the number of employees in the interval 1–9, Size 2 is 10–24, Size 3 is 25–49, Size 4 is 50–249, Size 5 is 250–499, and, finally, Size 6 is 500 and above.[6] Education is represented by 6 dummy variables as follows: no degree, primary school, secondary school, high school, vocational high school, and college & above. We control for workers' age as a quadratic polynomial. We also construct dummy variables for gender, marital status, full-time/part-time job status, and permanency status of the job. A specific feature of the Turkish labor market is that a non-negligible fraction of workers (around 24 percent) are employed informally. Information on both the formal and informal employment are nationally representative.

Table (1) reports the summary statistics for formal and informal employment separately. In our unweighted sample, we have 581,754 observations—441,866 of them are employed in formal jobs and 139,888 of them are employed informally. Relevant frequency weights are used to construct the summary statistics and in the regressions. The majority of the informal workers (more than 70 percent) are employed by size-1 firms, while only around 1.5 percent of them are employed by firms with 250 workers or more. A more balanced distribution is observed for formal employment. There is also a significant asymmetry between formal and informal workers in terms of educational attainment. Around 80 percent of informal workers have at most a secondary degree, while only 4 percent of them have college degree or above. For formal workers, however, 42 percent have at most a secondary degree, while around 29 percent have college degree or above.

Summary statistics for 87 industry categories, 27 occupation categories, and 26 geographical regions are not reported in Table (1). However, it will be useful to summarize the differences in industrial and occupational compositions across informal and formal workers. We start with industries. Around 10.7 percent of the informal workers are employed as salaried workers by firms in building construction, 10.5 percent in retail trade, 9 percent in agriculture, 8.7 percent in manufacturing of clothing, 8.4 percent in food and beverage services, and 5.3 percent in

[6]Note that we only know the number of workers employed by the firm, in which the survey subject is employed. We do not know whether the "size" measure reflects the number of full-time equivalent workers or not. Unfortunately, the dataset does not offer additional firm-level details.

road transportation. For formal workers, around 12.5 percent are employed by firms in public administration and defense, 9.5 percent in education, 8.3 percent in retail trade, 5.2 percent in health services, and 4.2 percent in manufacturing of clothing. These numbers suggest that informal workers are mostly employed in low-wage sectors. The occupational composition also draws a similar picture. For informal workers, 14.3 percent are in personal/security services, 9.8 percent are unskilled general workers in various sectors, 9.2 are unskilled sales/services personnel, 9.1 percent are in mine extraction/construction, 9 percent are in other craftsmanship, and 8.3 percent are general sales personnel. For formal workers, 9.5 percent are office clerks, 8.9 percent are in personal/security services, 7.7 percent are unskilled sales/services personnel, 7.2 percent are education professionals, 7.1 percent are machinery operators/assemblers, and 6.3 percent are drivers/operators. Clearly, both the industrial and occupational compositions state that informal workers are employed in sectors and jobs with relatively low skill requirements.

The difference between formal and informal hourly average real wages suggests that the unconditional gap is around 50 percent, which is prominent. The gap for monthly real wages is around 43 percent. This means that the informal workers not only get paid lower than formal workers, but also they work longer hours per week.

2.2 Institutional Background

In this section, we briefly discuss the institutional details about formal and informal jobs in Turkey to familiarize the reader with the related concepts and issues. The dataset that we work with allows us to see whether the worker is registered for social security or not. If the worker is not registered, we mark the worker as "informal" and the registered workers are marked as "formal." We would like to make it clear at this stage that we do not observe whether the firm, in which the worker is employed, is itself an informal firm or not. There are also other, rather obscure, forms of informality in Turkey. For example, some workers are registered for social security, but they agree with the employer to declare lower wages to reduce the tax burden on both parties. Our dataset does not allow us to track this type of informal employment. To

summarize, our definition of informality relies on the survey subjects' responses on whether they are registered for social security or not. As the summary statistics suggest, around 24 percent of all workers are employed informally based on this definition.

There are strong incentives for the firms to employ informal workers in the Turkish labor markets. The tax burden is high and enforcement is not very strong.[7] The high tax burden is often regarded as an obstacle in front of improving employment outcomes in Turkey. Based on this view, several subsidy programs have been implemented to support the employment prospects of disadvantaged groups after the onset of the 2008 crisis [see, e.g., Balkan, Baskaya, and Tumen (2014)]. Several regional subsidy programs have also been implemented to improve employment conditions in less-developed regions [see, e.g., Betcherman, Daysal, and Carmen (2010)]. The subsidies are generally in the form of reducing the employers' tax burden for new hires. There is also a possibility that the informal worker to declare herself as unemployed for the purpose of benefiting from unemployment compensation, which is offered for only a limited time. But, there is little practical scope for this possibility, since eligibility conditions of unemployment compensation require that the worker has recently left a formal job and has been actively seeking for a new one.

The pool of informal workers consists only of Turkish citizens. Although there might also be informally-employed foreign immigrants, the THLFS does not interview these people. This means that foreign immigrants have been ignored in constructing Turkish labor force statistics. This was not a bad approximation until mid-2012, because the share of these workers was negligible. However, after the onset of the civil war in Syria in mid-2012, Turkey has started to receive immigrant inflows from the Southern border and the share of foreign immigrants has probably become non-negligible. Our dataset covers the period 2006–2012, so the analysis is not affected by the Syria effect.

There is potentially a significant degree of segmentation along the formal-informal divide in the Turkish labor markets. In other words, the fundamentals driving the market forces are

[7] It is worthwhile to note that there are serious recent efforts to strengthen enforcement through more intensive policing and computerized monitoring.

potentially different across formal and informal jobs. There are several papers examining the degree of segmentation in the Turkish labor markets [see, e.g., Tansel and Kan (2012)]. Although there is no visible obstacle preventing the transitions from informal employment to formal employment, we see that there are substantial differences between the pool of workers employed in formal and informal jobs [see Table (1)]. There are several reasons suggesting that formal/informal segmentation is a realistic assumption for the labor markets in Turkey. First, the social security system in Turkey provides reasonable health and social insurance arrangements even for those who do not have any kind of official social security coverage. There are a lot of individuals aligned on this margin, who are willing to accept higher informal wages in exchange for better coverage in a lower-paying formal job. Second, the coverage provided by formal jobs is so generous that the spouses—in particular, wives who do not have a social security registration through a formal job—are also fully covered. This provides incentives for these spouses to work in informal jobs under more flexible terms. They are mostly not interested in the formal job market. Finally, firms also have incentives to create informal employment due to cost considerations. Taxes, social security contributions, and job security arrangements sum up to an important amount and can induce firms, especially the small ones, to search for workers in the informal market.

2.3 Econometric Model and Results

We run a least squares regression of log hourly wages on a comprehensive set of control variables including gender, marital status, age (as a quadratic), education, firm size, full-time/part-time work status, permanency of the job as well as year, region, industry, and occupation fixed effects. So, we control for a wide range of individual-level, group-level, and job-specific characteristics. Separate regressions are estimated for formal and informal employment. Specifically, we use the following equation to estimate firm-size wage gaps for formal and informal jobs:

$$\ln(w_i) = \beta_1 + \beta_2 s_2 + \beta_3 s_3 + \beta_4 s_4 + \beta_5 s_5 + \beta_6 s_6 + \boldsymbol{\theta}' \boldsymbol{X}_i + \epsilon_i, \tag{2.1}$$

Dependent variable: Natural logarithm of hourly real wages				
	Informal		**Formal**	
Covariate	**Coefficient**	**(SE)**	**Coefficient**	**(SE)**
Size 1	omitted		omitted	
Size 2	0.128***	(0.0004)	0.091***	(0.0002)
Size 3	0.168***	(0.0004)	0.105***	(0.0002)
Size 4	0.233***	(0.0005)	0.157***	(0.0002)
Size 5	0.301***	(0.0014)	0.216***	(0.0002)
Size 6	0.383***	(0.0021)	0.287***	(0.0002)
Female	-0.170***	(0.0003)	-0.077***	(0.0001)
Married	0.060***	(0.0003)	0.091***	(0.0001)
Full-time Job	0.439***	(0.0007)	0.392***	(0.0005)
Permanent Job	0.095***	(0.0004)	0.165***	(0.0004)
Primary School	0.016***	(0.0004)	0.026	(0.0004)
Secondary School	0.013***	(0.0004)	0.136***	(0.0004)
High School	0.130***	(0.0005)	0.278***	(0.0004)
Vocational High School	0.125***	(0.0005)	0.302***	(0.0004)
College & Above	0.463***	(0.0009)	0.709***	(0.0004)
Age	0.044***	(0.0001)	0.048***	(0.0001)
$\text{Age}^2/100$	-0.049***	(0.0001)	-0.046***	(0.0001)
Year Dummies	Yes		Yes	
Region Dummies	Yes		Yes	
Industry Dummies	Yes		Yes	
Occupation Dummies	Yes		Yes	
Constant	0.182***	(0.0013)	0.051***	(0.0012)
# of Observations	139,888		441,866	
R^2	0.50		0.58	

Table 2: **Estimation Results.** Size 1, male, non-married, part-time, temporary job, and no degree categories are the ignored dummy variables; so, the coefficients should be interpreted relative to these categories. Standard errors, reported in parentheses, are robust to heteroskedasticity and serial correlation. Appropriate frequency weights are used.

where s_1–s_6 denote the dummy variables for the six firm-size categories (note that size 1 is the omitted category), $\boldsymbol{X}_i$ is a vector of observed covariates, and ϵ_i is an error term.

Table (2) presents our estimates. According to our estimates, the wage gap between smallest and largest firms is 33.2 percent $[(\exp(0.287) - 1) \times 100]$ for formal employment, whereas it goes up to 46.7 percent $[(\exp(0.383) - 1) \times 100]$ for informal employment.[8] Moreover, this difference increases monotonically as size increases from 1 to 6. Specifically, when we consider wages as a function of size, our estimates suggest that the slope of this function is steeper for informal jobs than that for formal jobs. Figure (2.1) provides a plot of our estimates for formal versus informal employment. The upper panel is for the coefficients and the lower panel is for the gap itself. It is clear that the firm-size wage gap has a different slope along the

[8]See, e.g., Halvorsen and Palmquist (1980) about the rules of conversion for the regression coefficients.

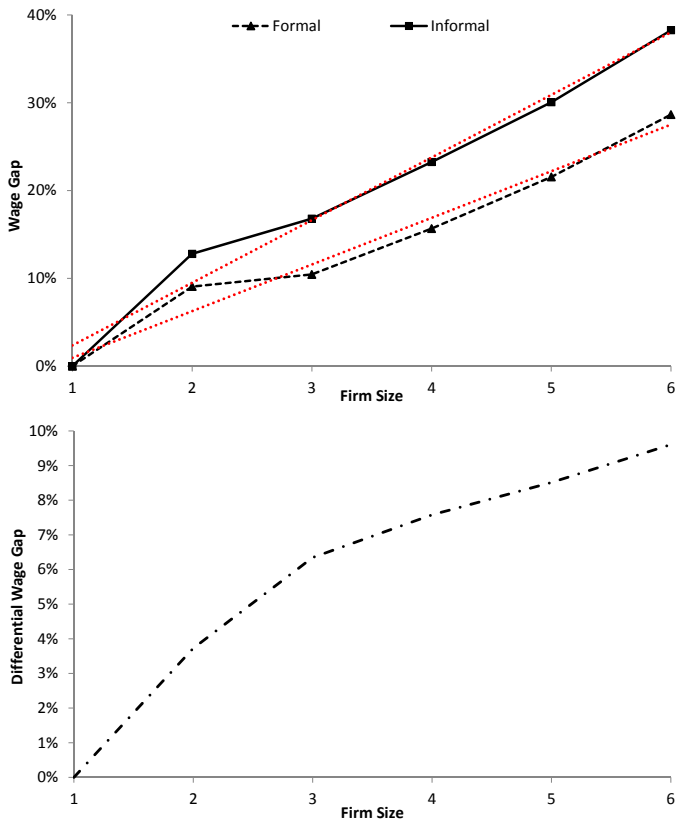

Figure 2.1: Plot of the estimates. The upper panel describes the firm-size wage gap relative to Size 1 for both formal and informal jobs. For example, a typical formal job at a Size 4 firm pays 17 percent $[(\exp(0.157) - 1) \times 100]$ more than a formal job at a Size 1 firm, while this gap is 26.2 percent $[(\exp(0.233) - 1) \times 100]$ for a typical informal job at a Size 4 firm. The red dashed lines are simply reference lines indicating the slope differential between the two black lines. To make this differential more concrete, the lower panel plots the vertical distance between the two lines in the upper panel. See Table (1) for the exact numbers used to construct the plots.

formal-informal divide. To be specific, the size wage gap is larger for informal employment than formal employment. To demonstrate this finding more clearly, we plot red dashed lines, which are just simple trend lines indicating that the wages increase faster with respect to size for informal employment than it does for formal employment.[9] It will perhaps be useful at this stage to mention the following point; the figure may mislead the reader in the sense that it might suggest that the formal-informal wage gap is zero at size 1 category. This is not the case. Since size 1 is the omitted category in our regressions, the results should be read "relative to size 1" for both formal and informal employment.

[9]Note also that the constant term is larger for the informal-employment regression than that for the formal-employment one. If the opposite case were true, then one would suspect that a lower intercept term for the informal-employment regression was partly responsible for the patterns of formal-informal size wage gap differentials we report.

Section 3 provides theoretical explanations for this empirical observation and discusses the underlying economic forces. Before switching to the theoretical model, there is a need to discuss the potential statistical problems that might have contaminated the OLS estimates reported in Table (2). The standard argument in the literature is that workers generally sort into firms of different sizes on a non-random basis; thus, the OLS estimates might potentially suffer from selectivity bias [Idson and Feaster (1990)]. The estimates might have been plagued with both "selection on observables" and "selection on unobservables" problems. Typically, these problems are resolved by implementing propensity score matching (PSM), Heckman selection-correction, or fixed-effect panel data techniques—conditional on the availability of appropriate data. A second potential bias might result from ignoring firm-level unobserved heterogeneity. Larger firms are potentially more productive and, given the quality of the labor input, the jobs that large firms open may produce more output than those that small firms open [Idson and Oi (1999)]. The main problem in addressing this issue is that firm-level details are generally not provided in large survey-based labor force datasets and controlling for firm-level unobserved (and even observed) heterogeneity is not a straightforward task. Third, it might be the case that the worker- and firm-level variables are subject to systematic measurement errors [Troske (1999)]. Standard error-correction techniques might be implemented to remove these errors. Finally, there might be behavioral factors—such as differences in organizational behavior etc.—across the firms of different sizes that could bias the results, unless they are appropriately controlled for [Idson (1996)]. In such a case, reduced-form parameters would not reflect the true impact of firm size on wages; instead, structural microeconometric models would be required to secure identification.[10]

Based on these ideas, we perform additional exercises to check the robustness of our empirical results. Note that the robustness exercises that we can apply are constrained by the availability of appropriate data. The main potential problem with the estimates reported in Table (2) is that workers employed in firms of different sizes may not be comparable in terms of both

[10]Other methods are also used in the literature to deal with the endogeneity problem. Brunello and Colussi (1998) apply a two-step procedure to remove endogeneity issues and find that, for Italy, the size premium is ignorable—although they state that their results are remarkably sensitive to the specification of the correction exercise. Muravyev (2009) focuses on workers who simultaneously hold two jobs in firms of different size and achieve identification by differencing across two jobs.

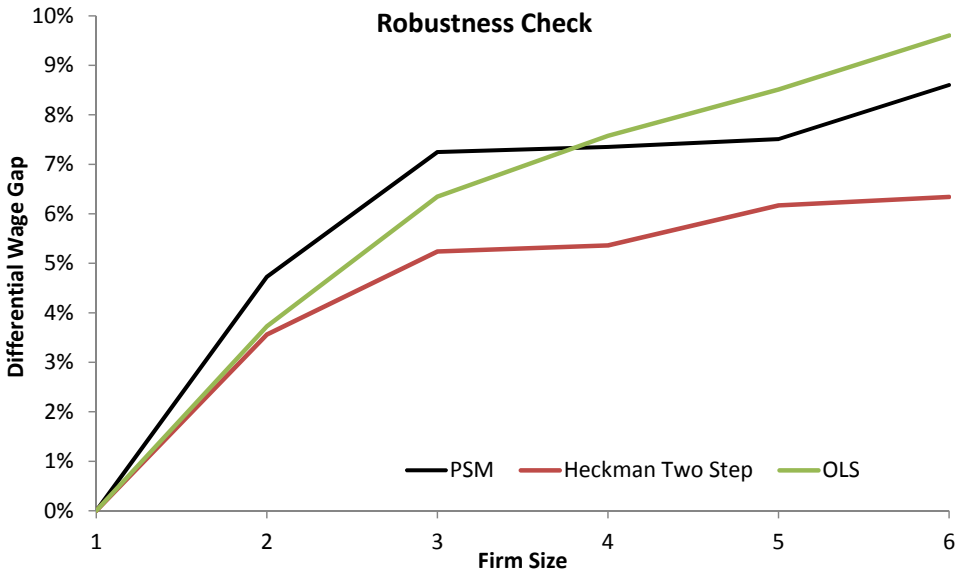

Figure 2.2: Robustness check. This figure compares, for robustness purposes, the OLS estimates (green) with the estimates obtained from PSM (black) and the Heckman two-step method (red).

observables and unobservables.[11] In other words, the estimates might be contaminated by both "selection on observables" and "selection on unobservables" problems. We perform basic checks of robustness to roughly test whether our estimates are plagued with these problem. To deal with selection on observables, we use the PSM method [see, e.g., Rosenbaum and Rubin (1983)]. We choose the Size 1 firm as the control group and assign the rest of the size categories one-by-one into the treatment category. This procedure is followed for formal and informal employment separately. The set of regressors used to derive the propensity score for each worker is the same as the set of regressors used in the OLS estimates reported in Table (2). To deal with selection on unobservables, we employ the standard Heckman two-step selection-correction method [see, e.g., Heckman (1979)]. The first step is a firm-size regression, while the second-step is for wages.[12] Note that our dataset does not offer a natural exclusion restriction, which leads us to use the same set of regressors in both steps. Although econometric identification in this setup may rest on functional forms, the estimates still give us an idea about the direction of the formal-informal differences in the size wage gap.

[11] Since the dataset we use offers a rather limited number of control variables for human capital and working conditions, these problems might indeed contaminate the estimates. See, for example, Fenn and Ashby (2004) for a study using a much larger set of control variables in an empirical exercise related to establishment size.

[12] The binary choice at the first step fixes Size 1 and assigns the rest of the size categories one-by-one as the treatment category. The inverse Mills ratios are inserted into the wage equations and the outcomes are predicted for formal and informal employment separately. This procedure is repeated 5 times to compare the Size 1 outcomes with the outcomes for all the other sizes.

The robustness analysis provides suggestive evidence that the qualitative nature of the OLS estimates is robust; that is, the size wage gap is more pronounced for informal employment than formal employment [see Figure (2.2)]. The magnitude of the gap, however, might change depending on the setup. It should perhaps be noted that the lack of firm-level details limits our ability to check the influence of firm-level heterogeneity issues on our estimates.

3 Theoretical Framework

In this section, we construct a theoretical model to explain/justify the empirical facts documented in the previous section. Our main purpose in this section is to theoretically identify the factors that can potentially lead to a larger size wage gap for informal employment than formal employment. We first develop a general framework that will serve as a benchmark model in our theoretical analysis. Then, we show how one can use this benchmark framework in understanding the differences in the observed firm-size wage gap patterns along the formal-informal divide. To be precise, we identify two potential channels. The first one argues that these differences can be attributed to government policies; that is, high taxes on formal employment may be generating a wedge between size wage gaps in the formal and informal jobs. This result implies a more general hypothesis: there is a negative relationship between the firm-size wage gap and the employment taxes. We use state-level microeconomic data from the US to test the validity of this "implied" hypothesis. Second, we show that relatively lower demand for high-pay informal jobs (than high-pay formal jobs) can lead to a higher size wage gap for informal employment than formal employment.

3.1 A Wage-Posting Game

This sub-section describes the benchmark model without making a distinction between formal and informal employment. The formal-informal divide will be introduced in the next sub-section. The model draws on the simple wage-posting game developed by Montgomery (1991), Lang (1991), and Lang, Manove, and Dickens (2005). Workers are homogeneous, i.e., they are equally productive; so, firms do not make any distinction/discrimination among them. Firms,

on the other hand, are heterogeneous. These assumptions are consistent with the empirical analysis conducted in Section 2. Our regressions control for all observed worker characteristics, but we do not have much information about firm characteristics. All we know is the size of the firm that the worker is employed.[13] There is an extensive literature empirically documenting the fact that firm size and productivity are strongly positively correlated. For example, Simon and Bonini (1958), Axtell (2001), and Luttmer (2007) show that the firm-size distribution is of the Pareto form, which suggests that firms on the right tail (i.e., larger firms) are scarce. This scarcity is due to the fact that they are, on average, more productive than the smaller firms.[14] Let N denote the size of a firm and z is the productivity. We assume that (1) firms differ in their sizes N and (2) productivity z and size N are related via a continuously increasing and invertible function g, i.e., $z = g(N)$ with $g'(\cdot) > 0$.

Each firm has one vacant position. Vacancies come with a posted wage. Workers have perfect information on all posted wage offers and, given this menu of wages, they choose which firm to apply. Each worker can apply for only one position. Workers know that higher wage offers will attract more applicants, which means that the probability of getting accepted will be smaller when the posted wage is high. Vacancy creation and wage posting are simultaneous events; thus, firms do not know the exact number of applicants when they choose a posted wage offer. They, instead, form expectations on the number of applicants. Firms hope to attract at least one applicant, because not being able to fill the position will be costly. In forming expectations on the number of applicants they will receive, the firms act on the information that higher wage offers will increase the expected number of applicants, which means that the probability of ending up with an unfilled vacancy will be lower.

The equilibrium will be analyzed within a two-stage game in this model. At the first stage, firms simultaneously post wage offers. At the second stage, workers observe all of the posted

[13] "Size" of a firm refers to the total number of workers employed by the firm.

[14] Note, however, that whether the empirical link between size and productivity features a causal relationship or not is a controversial issue. For example, Lallemand, Plasman, and Rycx (2005) document that the evidence is often indirect. It should also be noted that there is a huge theoretical literature featuring assortative matching of more able workers to more productive firms à la Becker (1973). Through such a mechanism, better managers/CEOs are employed in more productive and larger firms [see, e.g., Tervio (2008) and Gabaix and Landier (2008)]. However, this mechanism does not claim at all that productivity causes size or vice versa.

wages and they simultaneously decide which job to apply. Firms' wage posting strategy will be a best response to the expected worker behavior and workers' application strategy will be a best response to the observed wage offers. The resulting equilibrium will be *subgame perfect*. However, this will not be a standard subgame-perfect Nash equilibrium. The reason is that, in Nash equilibrium, agents know that their own actions will generate a response in market prices. In the present model, workers are small (i.e., they are price takers) and their actions will not affect market prices in any sense, because firms will act on "expected" number of applications rather than the actual numbers. Thus, following Lang, Manove, and Dickens (2005), we call our equilibrium a *subgame-perfect competitive equilibrium*.

There is a large (finite) number L of firms and the total number of job applicants in the job market is a random variable X, with mean μ_x. The realization of X is not observed by the firm; however, the firm knows and acts on μ_x. We assume that applicants come from a large population of workers, who make independent and equally probable decisions to enter the labor market. This definition implies that the total number of applicants, X, is Poisson distributed. We also assume that firms are able to commit to their equilibrium strategies, i.e., they will not change their posted wages after seeing the realization of X. In this setting, the firm will not hire anyone if it receives no applications; it will hire one worker at random if it receives more than one applications; and it will hire the sole applicant if it receives one and only one application.

Let j index the hiring firms. In this setup, posted wage offers can be represented by a vector $\boldsymbol{w}$ with firm-specific entries, w_j. Workers observe the wage profile $\boldsymbol{w}$ and develop a mixed strategy $\boldsymbol{h}(\boldsymbol{w})$, taken the wage offers as given. $\boldsymbol{h}(\boldsymbol{w})$ is a vector of application probabilities with entries $h_j(\boldsymbol{w})$, which describes the probability of applying to firm j given the entire wage offer profile $\boldsymbol{w}$. Workers are identical, so the strategy that they adopt is symmetric. As a result, the number of job applicants for any given firm j is drawn from a Poisson distribution with mean μ_j, which can be formulated as

$$\mu_j = h_j(\boldsymbol{w})\mu_x. \tag{3.1}$$

In other words, μ_j is the number of job applicants that firm j expects or hopes to attract, while μ_x is the expected number of total applicants in the job market. Then, firm j's expected profits can be expressed as

$$\pi_j = (1 - e^{-\mu_j})[g(N_j) - w_j], \tag{3.2}$$

where $g(N_j)$ is the value of a filled job's output as a function of firm size and $1 - e^{-\mu_j}$ is the probability that the vacancy is filled.[15] Now we are ready to describe the worker behavior, firm behavior, and the resulting subgame-perfect competitive equilibrium.

3.1.1 Job Application Strategy

In this subsection, we will describe the unique symmetric equilibrium in the worker application subgame. Let $f(\mu_j)$ describe the hiring strategy of firm j, who expects to receive μ_j applications with a posted wage w_j. More specifically, $f(\mu_j)$ is the probability that an additional designated applicant will be hired by firm j. We will describe how $f(\mu_j)$ is formulated in Section 3.1.2. Based on this definition, the expected income that any worker will receive by applying to firm j, which we denote with M_j, is simply

$$M_j = w_j f(\mu_j). \tag{3.3}$$

Given that all workers observe the entire profile of wage offers $\boldsymbol{w}$ and they observe $f(\mu_j)$'s, they construct a menu (or vector) $\boldsymbol{M}$ of expected incomes from all applications. Let $M = \max_j M_j$ is the maximum of all expected incomes. Workers will only apply to those firms with $M_j = M$. Following the terminology introduced by Lang, Manove, and Dickens (2005), we call M the "market expected income." So, in any symmetric equilibrium for the worker-application

[15] As a property of the Poisson distribution, $e^{-\mu_j}$ is the probability that the firm receives no applications, given that the Poisson arrival rate is μ_j. Then $1 - e^{-\mu_j}$ is the probability that firm j receives at least one application. By the firm's hiring strategy described above, we know that the firm will definitely choose to hire if it receives at least one application. As a result, $1 - e^{-\mu_j}$ is the probability that the vacancy is filled.

subgame, we have

$$M_j = \begin{cases} M, & \text{for } w_j \geq M, \\ w_j, & \text{for } w_j < M, \end{cases} \tag{3.4}$$

which suggests the following: (i) if the posted wage is less than the market expected income, then no worker will apply to such a firm; and (ii) if the posted wage offer is greater than or equal to the market expected income, then the firm will receive at least one application and the number of applications that the firm receives in the equilibrium will drive M_j down until $M_j = M$ is reached. This can be expressed as follows:

$$\begin{cases} \mu_j > 0, & \text{for } w_j \geq M, \\ \mu_j = 0, & \text{for } w_j < M. \end{cases} \tag{3.5}$$

Using (3.3) and (3.5), and assuming that f is invertible, it is possible to solve for μ_j, when $w_j \geq M$, as follows:

$$\mu_j = f^{-1}(M/w_j). \tag{3.6}$$

The total expected number of job applicants is

$$\sum_j \mu_j = \mu_x = \sum_{j|w_j \geq M} f^{-1}(M/w_j). \tag{3.7}$$

Equation (3.7) can be used to determine the equilibrium value of M. The left-hand size is a constant and the right-hand side is a decreasing function of M. As a result, we obtain a unique equilibrium solution for M, which we denote with $M^*(\boldsymbol{w})$. Then, $M^*(\boldsymbol{w})$, Equation (3.6), and Equation (3.1) jointly define the unique symmetric equilibrium $\boldsymbol{h}^*(\boldsymbol{w})$ in the worker application subgame with posted wage offers $\boldsymbol{w}$, given that firms are behaving optimally. This completes the characterization of the worker's application strategy.

3.1.2 Hiring Strategy

Let the profile of posted wages be $\boldsymbol{w}$. There is a potential pool of applicants for firm j. Since these applicants are identical, each of them has the same probability of applying to firm j. As we discuss above, μ_j is the number of workers that the firm expects to receive an application. Suppose now that an additional worker applies to firm j. The probability that this additional designated applicant will be hired is

$$f(\mu_j) = \sum_{k=0}^{\infty} \frac{1}{k+1} \frac{e^{-\mu_j}\mu_j^k}{k!}. \tag{3.8}$$

After some algebra, it is possible to represent this probability simply as

$$f(\mu_j) = \begin{cases} 1, & \text{for } \mu_j = 0, \\ (1 - e^{-\mu_j})/\mu_j, & \text{for } \mu_j > 0. \end{cases} \tag{3.9}$$

In words, when the expected number of applicants to firm j is zero, then the designated applicant will be hired with probability 1. If, on the other hand, the expected number of applicants is strictly greater than zero, then the probability that an additional applicant will be hired is strictly less than 1 and is a function of the expected number of applicants.

3.1.3 The Subgame-Perfect Competitive Equilibrium

The game between firms and workers yields a subgame-perfect competitive equilibrium described by the pair of behavioral profiles $\{\boldsymbol{w}^*, \boldsymbol{h}(\cdot)\}$. In this equilibrium, the mixed strategy $\boldsymbol{h}^*(\cdot)$ is symmetric across workers given a wage profile $\boldsymbol{w}^*$. From Equation (3.6), we know that $w_j = M^*(\boldsymbol{w})/f(\mu_j)$. Substituting this expression into firm's expected profit function given by Equation (3.2) and using firm's hiring strategy given in Equation (3.9), we find

$$\pi_j = (1 - e^{-\mu_j})g(N_j) - M^*(\boldsymbol{w})\mu_j. \tag{3.10}$$

The firm takes workers' application strategy $M^*(\boldsymbol{w})$ as given and maximizes the expected profits over the expected number of applicants μ_j. The first-order condition for this maximization

problem is simply

$$\mu_j = \log\left(\frac{g(N_j)}{M^*(\boldsymbol{w})}\right). \tag{3.11}$$

Manipulating Equation (3.11) yields the expression

$$M^*(\boldsymbol{w}) = g(N_j)e^{-\mu_j}. \tag{3.12}$$

The equilibrium operating profit of each firm then becomes

$$\pi_j = \left[1 - (1+\mu_j)e^{-\mu_j}\right] g(N_j) \tag{3.13}$$

and the equilibrium posted wage for each firm j becomes

$$w_j = \frac{g(N_j)\mu_j}{e^{\mu_j} - 1}. \tag{3.14}$$

Equation (3.14) is the core result in this section. It formulates posted wages as a function of two objects: (1) the value or productivity of a vacancy, $g(\cdot)$, which itself is a function of firm size, N_j and (2) the expected number of applicants to the position posted by firm j. This formula suggests that larger firms pay higher wages, because they are more productive and they expect to receive a larger number of applicants per vacant position.

To map this formulation into our empirical analysis, we differentiate wages with respect to size, which will give us how wage offers change as a response to an incremental increase in firm size. Note that the expected number of applicants will also be affected in this differential system. Thus, to get the full response, we totally differentiate Equation (3.14) with respect to w_j, N_j, and μ_j, which, after some algebra, gives the following expression:

$$\frac{dw_j}{dN_j} = \frac{g'(N_j)\mu_j}{e^{\mu_j} - 1} + \frac{g(N_j)}{e^{\mu_j} - 1}\left[1 - \frac{\mu_j e^{\mu_j}}{e^{\mu_j} - 1}\right]\frac{d\mu_j}{dN_j} > 0. \tag{3.15}$$

This formula can be interpreted as follows. There is a consensus in the literature that larger firms pay higher wages to observationally equivalent workers. This is our observation also for

Turkey. (See our estimates given in Table (2).) This fact is reflected above as $dw_j/dN_j > 0$. dw_j/dN_j has two components. The first component $g'(N_j)\mu_j/(e^{\mu_j} - 1) > 0$ says that larger firms pay higher wages because they are more productive. The sign of the second term is negative because the term $1 - (\mu_j e^{\mu_j})/(e^{\mu_j} - 1)$ is less than or equal to zero for $\mu_j \geq 0$. Based on the mechanism we describe, we also know that $d\mu_j/dN_j > 0$; that is, larger firms expect to attract a greater number of applicants, everything else constant. This greater pool of applicants, however, generates a secondary effect: the effect of size on wages will be mitigated by the fact that, with such a large applicant pool, the firm can easily fill its vacancy without the need to pay a large size premium.

To summarize, our model suggests that the effect of size on wages operates through two channels: the productivity effect and the labor supply effect. The productivity effect makes a positive contribution, while the labor supply effect makes a negative contribution to the magnitude of dw_j/dN_j. But, in the overall, empirical evidence suggests that the productivity effect dominates and the sign of dw_j/dN_j remains positive. In the next subsection, we will extend this mechanism by assuming segmented markets and, then, use this benchmark model to explain the stylized fact we document in Section 2.

3.2 Firm-Size Wage Gaps in Segmented Markets

From this point on, we assume that markets are segmented, i.e., formal and informal jobs are posted in separate markets. This assumption is empirically justified and is often invoked in theoretical work. In particular, influential papers including Dickens and Lang (1985) and Heckman and Hotz (1986) document that segmented markets along the formal/informal divide exist in the real world. Following this tradition, we assume that there is such a segmentation in our theoretical environment. This means that the supply and demand conditions underlying the equilibrium solution are different for formal and informal jobs. See Section 2.2 for additional discussion about the validity of the segmentation assumption in the Turkish labor markets.

In the rest of this section, we present the detailed predictions of the benchmark model devel-

oped above by assuming that labor markets are segmented along the formal/informal divide, firms post separate job openings in formal and informal job markets, and workers' preferences toward jobs also support this segmentation. The predictions of the model are twofold. First, the model predicts (as we discuss in Section 3.2.1) that setting high tax rates on formal employment inserts a wedge between the firm-size wage gaps for formal and informal jobs. Second, and finally, high-paying informal jobs attract a smaller number of applicants than the high-paying formal jobs do. We discuss below in greater detail the theoretical predictions, our interpretations of them, and additional/auxiliary empirical evidence to check the validity of those predictions.

3.2.1 The Role of Taxes

An important difference between formal and informal jobs is that formal jobs are more costly tax-wise for both the employer and employee. For this reason, informal jobs are often associated with tax avoidance behavior. Taxes are dropped from the formulas in the previous subsection for expositional simplicity. Below we make the tax rate visible and discuss how taxes affect behavior in our model.

Suppose formal and informal jobs are offered in segmented markets, but the only difference between these two markets is that those working in formal jobs pay taxes while the informal workers do not. This suggests that the Equation (3.14) becomes

$$w_j = \frac{g(N_j)\mu_j}{(1+\tau)(e^{\mu_j}-1)}, \tag{3.16}$$

where $\tau \in (0,1)$ is the tax burden on the formal job. The informal job sets $\tau = 0$, while for the formal job $\tau > 0$. So, it is also possible to re-label formal and informal jobs as high-tax and low-tax jobs, respectively. Using (3.15) and (3.16), it is clear that

$$\frac{d^2 w_j}{dN_j d\tau} < 0, \tag{3.17}$$

which means that the size gap is a decreasing function of taxes. In other words, the size-

gradient of wages becomes smaller as the tax rate goes up. This formulation predicts that the size-wage gap should be higher for informal jobs than formal jobs. Taxes impose a wedge between the firm-size wage gaps in formal and informal jobs.[16] However, this prediction does not only apply to jobs along the formal/informal divide and it can be translated into a more general one: firm-size wage gaps is a decreasing function of employment taxes. So, if we have, say, cross-regional data on firm-size wage gaps and employment taxes, then our prediction suggests that we should observe lower firm-size wage gaps in low-tax regions. Next we formally test this prediction.

To be specific, we convert our theoretical prediction into an empirically testable one by making the following statement: if the tax rate is a relevant factor affecting the firm-size wage gaps, then it has to be the case that the firm-size wage gap should be lower in regions with higher taxes on labor. We perform this test using data from the United States by utilizing the labor tax differentials across the states.[17]

We use the Current Population Survey (CPS) March supplements from 2010 and 2011 to estimate the firm-size wage gaps for each state in our data set. We regress hourly log wages for full-time full-year workers employed by the firms in the private sector on gender, education, race, age (quadratic), marital status, occupation, industry, and firm size. The firm size is defined by a dummy variable taking 1 if the worker is employed by a firm with greater than 500 employees and taking 0 otherwise. We perform a separate regression for each of the states in our sample. To avoid the possible spurious size effects coming from unusual industry concentrations in smaller states, we focus on the 15 largest states in terms of their Gross Domestic Products. These states are California, Florida, Georgia, Illinois, Maryland, Massachusetts,

[16] This wedge does not correspond to a parallel shift in the wage-size relationship in the formal versus informal jobs. The tax rates (and the cost of employment in percentage terms) are higher for high-wage and/or high-rank jobs, which are mostly offered by larger firms. This is exactly in line with the results of our empirical and theoretical analyses.

[17] Ideally, one would like to see similar estimates for Turkey. However, since Turkey has uniform labor taxes over the country, it is not possible to extract a cross-sectional variation in labor taxes to test the link between taxes and size premia. Variation in taxes can also be captured in a cross-country analysis. There are some studies in the literature comparing the cross-country patterns in firm-size wage gaps. For example, Albaek, Arai, Asplund, Barth, and Strojer Madsen (1998) find that the size-elasticity of wages in Nordic countries—Denmark, Finland, Norway, and Sweden—is of almost the same order of magnitude when compared with the size-elasticity in the US. The labor markets in the US are subject to very different institutional rules and regulations (including taxes and wage setting practices) than the labor markets in the Nordic countries. Although this may suggest that the link between taxes and size premia seems weak, the cross-country comparisons reported in the literature are crude and lack an appropriate control of country fixed effects.

Michigan, New Jersey, New York, North Carolina, Ohio, Pennsylvania, Texas, Virginia, and Washington.

The Federal taxes are fixed, but the state-level labor taxes exhibit significant variation across the states. The state-specific tax burden is calculated from the state-level labor tax information provided by the US Department of Labor, Employment, and Training Administration. We multiply the maximum applicable salary reported with the corresponding tax rates to calculate the "maximum possible burdens" on employers in each state. This maximum burden is usually the regular burden for most of the states because the maximum applicable salary is fairly low for the average full-time full-year worker in the US.

We find a clear negative relationship between the labor tax rates and the firm-size wage gaps at the state level in the US. Figure (3.1) plots the state-level tax burden against the firm-size wage gap estimates. The negative relationship is even sharper when we rule out the outliers that emerge in our regressions (Florida, Massachusetts, and Washington). This result roughly confirms the validity of our theoretical predictions. We interpret this finding as follows. Each extra dollar offered by the firm is translated into less-than-a-dollar increase—due to taxes—in the wage offer received by the worker. This generates a downward pressure on the size premium in high-tax environments, while this pressure gets smaller as the tax rates decline. The informal job is the extreme case. As tax rates converge to zero, the firm-size wage gaps start increasing and, at the limit, it resembles the size premium in the informal job.

As we discuss in Section 2.3, the estimates reported in Figure (3.1) might suffer from various biases.[18] However, similar to the THLFS, the structure of the CPS dataset limits our ability to deal with all of these statistical problems. The March CPS is a cross-sectional dataset with yearly frequency and does not allow following individuals over time. Moreover, like all the other large cross-sectional datasets, it does not offer natural instrumental variables to be used in removing problems related to endogeneity issues. That said, there are still two

[18] See Section 2.3 for a discussion of these potential biases.

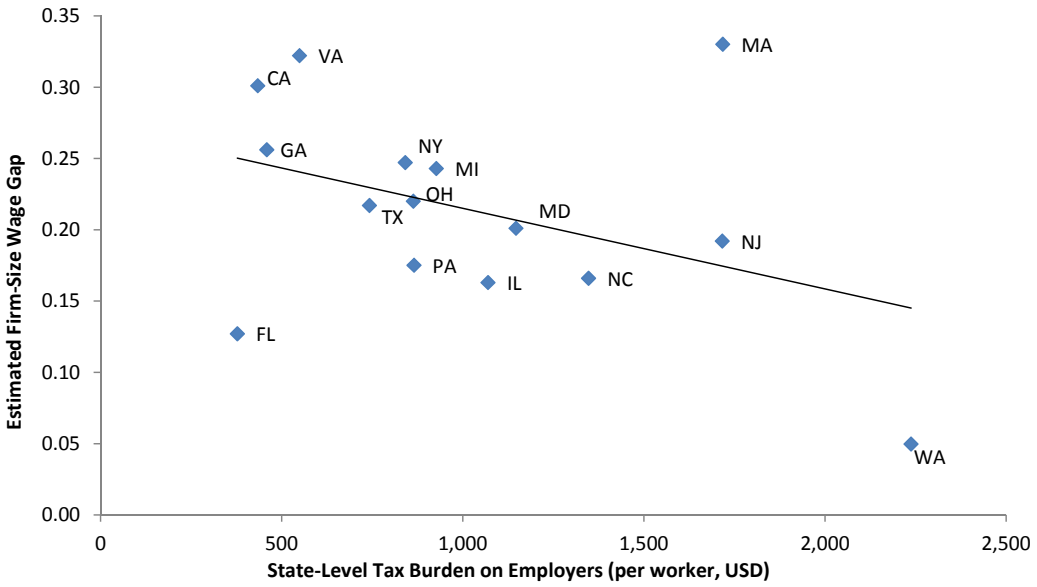

Figure 3.1: Firm-Size Wage Gap and Taxes.

alternative methods that we can use to check the robustness of the results reported above: propensity score matching and selectivity correction. In fact, these are the two methods that we use in Section 2.3 to test the reliability of our estimates for the Turkish data. We basically compare the OLS estimates with the estimates from the propensity score matching and selection-correction exercises. This comparison roughly informs us whether our estimates are contaminated by non-random sorting on observables and unobservables. For the propensity score matching exercise, we construct propensity scores using the same set of regressors we employed in our OLS estimates. For the selection-correction exercise, we do not use an exclusion restriction as the CPS does not offer a natural instrumental variable. We find that the negative relationship between firm-size wage gaps and taxes persists under these alternative specifications.[19] The magnitude of the negative relationship is aligned in descending order as follows: propensity score matching, OLS, and selection correction. This order suggests that selection on unobservables tends to generate an upward bias in the OLS estimates, while selection on observables has a potential to induce a downward bias. We would like to mention one final caveat. These estimates ignore two potentially relevant points: state fixed effects and firm-level details. These factors cannot be accounted for with the March CPS, since it does not have a panel dimension and does not report firm-level details.

[19] We do not report the estimates for the sake of brevity.

This finding suggests that labor-tax policy can affect how observationally equivalent workers fare in larger firms relative to smaller firms. In other words, we show that government policy can be a determinant of the magnitude of the firm-size wage gap and, therefore, the allocation of workers across firms of different sizes. Our results have a large perspective in terms of tax policy. It is well-established in the literature that decreasing the tax rates will likely reduce the size of informal employment and push the economic activity towards formality [see, e.g., Ihrig and Moe (2004)]. Since the size wage gap is higher for informal employment than formal employment, the magnitude of wage dispersion might decline as a response to decreased share of informality. However, not only the wage dispersion, but also the size distribution of firms might change in this case, because the incentives for staying small will be lessened when tax rates go down. As a result, there will be several general equilibrium effects that one needs to keep track of to have a more realistic sense of the policy impact of changing taxes on formal employment. Our model is too simplistic to reflect these effects. The ideal setup to fully address this issue would be an equilibrium search model, in which workers choose the sector they work, wage distribution is endogenous, and the size distribution of firms is also preferably endogenous [see, e.g., Meghir, Narita, and Robin (2014)]. After the parameters of such a model are structurally estimated, tax policy changes can be simulated to recover counterfactual outcomes.

There is one more—a rather provocative—policy implication, which is related to the ones discussed above. Our results communicate the message that informal employment generates large wage differentials. However, there is a serious identification issue. If the fraction of informal employment is disproportionately large and if informal employment is widespread in small firms, then the size distribution of firms itself may impose significant wage differentials. So it is unknown whether informality or the size distribution is driving the wage differentials. That the share of small firms—as well as the share of informal firms—shrinks along the development path is a well-established fact. Based on this observation, one may suggest that the developing countries should not focus on specific policies to reduce wage inequality; instead, the focus should be on economic development. Wage differentials would naturally become

smaller as the economy grows further and a greater fraction of large firms starts operating. That said, cross-country comparisons reveal that, in some developed countries (i.e., the United States), wage inequality can actually be quite large; thus, there might not be an unconditional relationship between development levels and wage differentials. Institutional forces, such as such as minimum wage laws, unionism, and unemployment compensation schemes, should be controlled for to have a better understanding of cross-country comparisons.

3.2.2 An Alternative Mechanism

In this part, we will show that our model can explain the stylized fact documented in Section 2 even when the tax dimension is shut down. Again, we assume that markets are segmented. Our starting point is the Equation (3.15), which is the main prediction of our benchmark model. The model predicts that two forces jointly generate the firm-size wage gaps: (i) larger firms pay higher wages since they are more productive (the productivity effect) and (ii) larger firms expect to attract a larger number of applicants and the size of the applicant pool determines the size premium (the labor supply effect). The latter is a mitigating force; that is, when the pool of potential applicants is large, the firm will tend to keep the size premium small.

Based on this prediction, the size gradient of equilibrium wages can be different between formal and informal jobs because of two reasons. First, it can be the case that formal jobs are more productive than informal jobs. In other words, given firm size, $g'(N)$ should be larger for formal jobs than informal jobs. However, we only have information about the formality status of the job, not the formality status of the firm. Informal jobs can be posted by formal firms and this is a common practice in Turkey. Without further evidence, asserting that formal jobs should be more productive than informal jobs may produce misleading results. Thus, we de-emphasize the potential differences in the productivity-effect channel.

The second reason is due to the potential differences in the size elasticity of the expected number of applicants. Larger firms offer higher wages (since they are more productive) than smaller firms. Higher wages, in turn, attract a larger number of applicants per job; thus, $d\mu_j/dN_j > 0$. However, the degree of this differential effect is potentially different between

formal and informal jobs. Let d_i and d_f denote the magnitude of this differential effect $d\mu_j/dN_j > 0$ for informal and formal jobs, respectively. The stylized fact we report—that is, the firm-size wage gap is larger for informal employment than formal employment—implies that $d_i < d_f$, so that dw_j/dN_j is larger for informal jobs than formal jobs. In other words, the implication that $d_i < d_f$ can itself explain why the firm-size wage gap is larger for informal employment.

What does $d_i < d_f$ say? It means that the mitigating effect—i.e., the labor supply effect, which operates as a negative force—is weaker for informal jobs. In other words, higher wage offers bring in a smaller number of additional applicants in informal jobs than formal jobs. So, the firm-size wage gap is larger for informal jobs even when the productivity effect is deactivated. Alternatively, it is also possible to argue that the wage penalty for working in a small firm is bigger for informal jobs than formal jobs, because the pool of applicants is larger for informal than formal low-wage jobs. Informal jobs are easier to dissolve, which suggests that the high negative wage penalty associated with informal jobs in small firms might be a consequence of an upward sloping labor supply curve to an individual employer that would be steeper for informal than formal jobs. This mechanism may be operating due to several reasons. First, larger firms, on average, may have greater skill requirements both for the formal and informal jobs they offer. However, informal jobs with higher skill requirements may fail to attract a large number of applicants, since skilled workers are expected to opt for formal jobs. Second, applicants may be valuing other job-specific (pecuniary and/or non-pecuniary) amenities along with pay. This kind of amenity packages are weaker in informal jobs, by definition. This weakness may itself lead to a smaller applicant pool in informal jobs, if the weight assigned to these side amenities are high enough.

Finally, information on job opportunities in informal jobs are most likely disseminated through informal job search networks (such as relatives, friends, neighbors, and other acquaintances) rather than formal channels (such as ads). This can potentially reduce the size of the applicant pool for informal jobs. This effect might be even stronger for large firms due to higher visibility. The Turkish HLFS dataset provides information about the search method by which

	Informal Jobs	Formal Jobs
Informal Search Methods	0.363	0.245
Formal Search Methods	0.637	0.755

Table 3: **Search Methods:** Turkish Household Labor Force Survey data between 2006–2012 are used in the analysis. Appropriate frequency weights are used.

the employed workers found their current jobs. Informal job search corresponds to job search through one's friends, relatives, and acquaintances. Formal job search methods include direct contact with the employer, sending CVs, filling online application forms, visiting public or private employment agencies, unions, graduate placement offices, etc. Table (3) provides information on the job search methods used to find informal versus formal jobs.

The data suggest that around 36.3 percent of all informal job holders have found their jobs through their social contacts, while this number goes down to 24.5 percent for formal job holders. In other words, the incidence of finding a job through informal connections is 50 percent higher for informal jobs than formal jobs. We conclude that the potential differences among the job search strategies of the workers seeking informal versus those seeking formal jobs may also be a relevant channel in explaining formal-informal differences in the size-wage gap patterns.

Other alternative explanations might also be put forward. For example, Weiss and Landau (1984) argue that size-wage differentials might arise also because of the greater quality of the workers in the queue of application for vacancies in large firms. Larger firms offer better human capital investment opportunities and, thus, a better career, which intensifies the competition for locating jobs in these firms. In other words, not only the quality of the existing workers, but also the quality of the potential workers in the queue matters to understand the size-wage differentials. In terms of our model, this mechanism suggests the following alternative explanation. If the average quality of the workers in the queue increases with size faster for formal jobs than informal jobs, this can also be a reason why the firm-size wage gap is larger for informal jobs than formal jobs.

Another alternative explanation might be related to size-tenure and size-education effects. To

be specific, if larger firms offer better career advancement opportunities to employees, then part of the size effect would reflect this. For example, if returns to education and returns to tenure are higher in large firms than those in small firms, then part of the size effect in large firms might have been absorbed by the coefficients of education and tenure variables. To see this more clearly, we perform our wage regressions conditioning on the size of the firm. We find that the returns to a college degree (relative to returns to no degree) is approximately 15 percentage points higher in Size-6 firms relative to Size-1 firms. We do not have a direct tenure variable to calculate the size-tenure effects. The age coefficient can capture the tenure effects only in a very rough way. Instead, our dataset features a variable for "tenure in last job." When we include this variable as a quadratic polynomial into the analysis and perform the same size-specific regressions, we indeed observe that returns to tenure in last job is around 0.049 percent in Size-6 firms, while it monotonically declines as firm size decreases and reaches 0.028 at Size-1 firms. Although these numbers do not carry causal meanings, they suggest that the formal-informal differences in firm-size wage gaps could be partly explained by size-education and size-tenure effects.

On a separate issue, El Badaoui, Strobl, and Walsh (2010) argue that, since larger firms are more likely to get caught if they operate informally, large high-wage firms will be in the formal sector and small low-wage firms will tend to be in the informal sector. Although this statement is about the formality status of the firms—not about the formality status of workers—and, in this sense, not quite similar to what we address in this paper, it still possible to relate our paper to theirs. Clearly, getting caught working informally will lead to job termination—plus fines. This suggests that informal jobs in large firms will be more risky than informal jobs in small firms, because large firms are more visible than small ones. This risk element can itself lead to larger size gradient for the wage function of informal workers than that for formal workers. As a result, similar to El Badaoui, Strobl, and Walsh (2010), our paper also features the idea that the enforcement dimension can also lead to different size effects for formal and informal jobs.

Finally, we would like to mention an interesting generality that our model can exhibit. Suppose

that there are two labor markets and these markets are segmented—which is the formal job market versus the informal job market in our context. Depending on the market structure, if higher wage offers in one market can bring in a lower number of applicants than the same offer can bring in the other market, then the size-wage gap is likely larger in the former than the latter. The finding presented in Fox (2009) hints that this conjecture is true. Specifically, using data from US and Sweden, Fox (2009) finds—and explains employing a different theoretical model—that the firm-size wage gap is larger for jobs with managerial responsibilities than the standard jobs. Notice that this is another example of segmented markets. Managerial and non-managerial jobs are traded in separate markets. If higher wage offers bring in a smaller number of additional applicants in managerial jobs than non-managerial jobs (which is most likely the case), then our theoretical framework predicts that the firm-size wage gap should be greater for jobs with managerial responsibilities than the non-managerial jobs. This confirms that our theoretical model applies to other environments as well.

4 Concluding Remarks

It is well-documented in the literature that observationally equivalent workers tend to receive higher wages in larger firms than smaller firms. In this paper, we report a new stylized fact using micro-level data from Turkey: the firm-size wage gap is larger for informal jobs than formal jobs. We develop an equilibrium wage-posting model to provide a systematic explanation for this fact. The model offers two alternative explanations. First, the model suggests that high tax burden on formal jobs creates a wedge between the size wage gaps for formal and informal jobs. We provide some auxiliary empirical support for this prediction using state-level tax differentials in the US. The second prediction of the model suggests that the size premium has two components. The first one is a pure productivity effect, i.e., larger firms offer higher wages because the jobs that they open are more productive than those opened by smaller firms. This effect is homogeneous across formal and informal jobs. The second one is a mitigating labor supply effects. It says that the initial wage offer will be somewhat lowered by the firm depending on the size of the applicant pool. If the pool is large,

then the mitigating effect can be strong. If, on the other hand, the applicant pool is small, then the mitigating labor supply effect is negligible. Our model predicts that the mitigating force is smaller for informal jobs relative to formal jobs. This, alone, can explain the fact we document. Although the fact that we document is quite robust to alternative specifications, further empirical research is needed to check the validity of the finding that the firm-size wage gaps are larger for informal jobs than formal jobs. This task can best be performed by analyzing micro-level datasets from other developing countries with a large share of informal jobs.

Other than documenting a new fact, this paper makes several contributions to the related literature. First, it communicates the idea that the firm-size wage gap has a heterogeneous nature across segments and/or sectors in the labor market. Second, it suggests that the firm-size wage gap is likely affected by government policy by means of taxes and this link is potentially more pronounced in developing countries, where informal employment is widespread. Finally, we argue that the pool of applicants in firms of different sizes may exhibit variation across formal and informal employment. We argue that this can be explained by various economic mechanisms ranging from job search networks to preferences for non-pecuniary job amenities.

Eliminating the informal economy by redesigning tax and enforcement policies has been a longstanding matter of interest for developing countries. Our analysis suggests that changing taxes on formal activity will have direct effects on the extent of aggregate wage dispersion in the labor market. The change in the wage dispersion will be driven by reallocations both within and between sectors. However, our model is too simplistic to reflect the effects of tax changes on the aggregate wage distribution in a realistic way. The ideal setup to address such a question would be constructing and structurally estimating an appropriately designed general equilibrium search model [see, e.g., Meghir, Narita, and Robin (2014)]. We leave this task for future research.

References

Albaek, K., M. Arai, R. Asplund, E. Barth, and E. Strojer Madsen (1998): "Measuring Wage Effects of Plant Size," *Labour Economics*, 5, 425–448.

Axtell, R. L. (2001): "Zipf Distribution of U.S. Firm Sizes," *Science*, 293, 1818–1820.

Baker, G. P., M. Jensen, and K. J. Murphy (1988): "Compensation and Incentives: Practice vs. Theory," *Journal of Finance*, 43, 593–615.

Balkan, B., Y. S. Baskaya, and S. Tumen (2014): "Evaluating the Impact of the Post-2008 Employment Subsidy Program in Turkey," Central Bank of the Republic of Turkey, WP#1414.

Barth, E. and H. Dale-Olsen (2011): "Employer Size or Skill Group Size Effect on Wages?" *Industrial and Labor Relations Review*, 64, 341–355.

Becker, G. S. (1973): "A Theory of Marriage: Part I," *Journal of Political Economy*, 81, 813–846.

Betcherman, G., N. M. Daysal, and P. Carmen (2010): "Do Employment Subsidies Work? Evidence from Regionally Targeted Subsidies in Turkey," *Labour Economics*, 17, 710–722.

Brown, C. and J. Medoff (1989): "The Employer Size-Wage Effect," *Journal of Political Economy*, 97, 1027–1057.

Brunello, G. and A. Colussi (1998): "The Employer Size-Wage Effect: Evidence from Italy," *Labour Economics*, 5, 217–230.

Dickens, W. T. and K. Lang (1985): "A Test of Dual Labor Market Theory," *American Economic Review*, 75, 792–805.

El Badaoui, E., E. Strobl, and F. Walsh (2010): "The Formal Sector Wage Premium and Firm Size," *Journal of Development Economics*, 91, 37–47.

Fakhfakh, F. and F. Fitzroy (2006): "Dynamic Monopsony: Evidence from a French Establishment Panel," *Economica*, 73, 533–545.

Fenn, P. and S. Ashby (2004): "Workplace Risk, Establishment Size and Union Density," *British Journal of Industrial Relations*, 42, 461–480.

Fox, J. T. (2009): "Firm-Size Wage Gaps, Job Responsibility, and Hierarchical Matching," *Journal of Labor Economics*, 27, 83–126.

Gabaix, X. and A. Landier (2008): "Why Has CEO Pay Increased So Much?" *Quarterly Journal of Economics*, 123, 49–100.

Garicano, L. and E. Rossi-Hansberg (2006): "Organization and Inequality in a Knowledge Economy," *Quarterly Journal of Economics*, 121, 1383–1435.

Gong, X. and A. van Soest (2002): "Wage Differentials and Mobility in the Urban Labour Market: A Panel Data Analysis for Mexico," *Labour Economics*, 9, 513–529.

Groshen, E. L. (1991): "Five Reasons Why Wages Vary Across Employers," *Industrial Relations*, 30, 350–381.

Halvorsen, R. and R. Palmquist (1980): "The Interpretation of Dummy Variables in Semilogarithmic Equations," *American Economic Review*, 70, 474–475.

Heckman, J. J. (1979): "Sample Selection Bias as a Specification Error," *Econometrica*, 47, 153–161.

Heckman, J. J. and J. V. Hotz (1986): "An Investigation of the Labor Market Earnings of Panamanian Males Evaluating the Sources of Inequality," *Journal of Human Resources*, 21, 507–542.

Hollister, M. N. (2004): "Does Firm Size Matter Anymore? The New Economy and Firm Size Wage Effects," *American Sociological Review*, 69, 659–676.

Idson, T. L. (1996): "Employer Size and Labor Turnover," *Research in Labor Economics*, 15, 273–304.

Idson, T. L. and D. J. Feaster (1990): "A Selectivity Model of Employer-Size Wage Differentials," *Journal of Labor Economics*, 8, 99–122.

Idson, T. L. and W. Y. Oi (1999): "Workers are More Productive in Large Firms," *American Economic Review*, 89, 104–108.

Ihrig, J. and K. S. Moe (2004): "Lurking in the Shadows: The Informal Sector and Government Policy," *Journal of Development Economics*, 73, 541–557.

Lallemand, T., R. Plasman, and F. Rycx (2005): "Why Do Large Firms Pay Higher Wages? Evidence from Matched Worker-Firm Data," *International Journal of Manpower*, 26, 705–723.

——— (2007): "The Establishment-Size Wage Premium: Evidence from European Countries," *Empirica*, 34, 427–451.

Lang, K. (1991): "Persistent Wage Dispersion and Involuntary Unemployment," *Quarterly Journal of Economics*, 106, 181–202.

Lang, K., M. Manove, and W. T. Dickens (2005): "Racial Discrimination in Labor Markets with Posted Wage Offers," *American Economic Review*, 95, 1327–1340.

Lazear, E. P. (2009): "Firm-Specific Human Capital: A Skill-Weights Approach," *Journal of Political Economy*, 117, 914–940.

Luttmer, E. G. J. (2007): "Selection, Growth, and the Size Distribution of Firms," *Quarterly Journal of Economics*, 122, 1103–1144.

Magnac, T. (1991): "Segmented or Competitive Labor Markets," *Econometrica*, 59, 165–187.

Manning, A. (2003): *Monopsony in Motion: Imperfect Competition in Labour Markets*, Princeton, NJ: Princeton University Press.

Marcouiller, D., V. Ruiz de Castilla, and C. Woodruff (1997): "Formal Measures of the Informal-Sector Wage Gap in Mexico, El Salvador, and Peru," *Economic Development and Cultural Change*, 45, 367–392.

Mazumdar, D. (1981): *The Urban Labor Market Income Distribution: A Study of Malaysia*, Oxford, UK: Oxford University Press.

Meagher, K. J. and H. Wilson (2004): "Different Firm Size Effects on Wages for Supervisors and Workers," *Economics Letters*, 84, 225–230.

Meghir, C., R. Narita, and J.-M. Robin (2014): "Wages and Informality in Developing Countries," Forthcoming, *American Economic Review*.

Mellow, W. (1982): "Employer Size and Wages," *Review of Economics and Statistics*, 64, 341–345.

Montgomery, J. D. (1991): "Equilibrium Wage Dispersion and Interindustry Wage Differentials," *Quarterly Journal of Economics*, 106, 163–179.

Mortensen, D. T. (2010): "Dynamic Monopsonistic Competition and Labor Market Equilibrium," Society for Economic Dynamics, 2010 Meeting Paper #164.

Muravyev, A. (2009): "Employer Size, Wages and Unobserved Skills: Evidence from Moonlighters in the UK," *Manchester School*, 77, 651–674.

Oi, W. Y. and T. L. Idson (1999): "Firm Size and Wages," in *Handbook of Labor Economics*, ed. by O. Ashenfelter and D. Card, New York, NY: Elsevier, vol. 3, chap. 33, 2165–2214.

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 Argentina? A Semiparametric Approach," *European Economic Review*, 50, 1817–1841.

Rosenbaum, P. R. and D. B. Rubin (1983): "The Central Role of the Propensity Score in Observational Studies for Causal Effects," *Biometrika*, 70, 41–55.

Simon, H. A. and C. P. Bonini (1958): "The Size Distribution of Business Firms," *American Economic Review*, 48, 607–617.

Stiglitz, J. E. (1976): "The Efficiency Wage Hypothesis, Surplus Labor, and the Distribution of Labour in LDCs," *Oxford Economic Papers*, 28, 185–207.

Tan, H. and G. Batra (1997): "Technology and Firm Size-Wage Differentials in Colombia, Mexico, and Taiwan (China)," *World Bank Economic Review*, 11, 59–83.

Tansel, A. (1999): "Formal versus Informal Sector Choice of Wage Earners and their Wages in Turkey," Economic Research Forum, WP#9927.

Tansel, A. and E. O. Kan (2012): "Labor Mobility across the Formal/Informal Divide in Turkey: Evidence from Individual Level Data," IZA Discussion Paper #6271.

Tervio, M. (2008): "The Difference That CEOs Make: An Assignment Model Approach," *American Economic Review*, 98, 642–668.

Troske, K. R. (1999): "Evidence on the Employer Size-Wage Premium from Worker-Establishment Matched Data," *Review of Economics and Statistics*, 81, 15–26.

Weiss, A. and H. J. Landau (1984): "Wages, Hiring Standards, and Firm Size," *Journal of Labor Economics*, 2, 477–499.