PEDL Research Papers

This research was partly or entirely supported by funding from the research initiative Private Enterprise Development in Low-Income Countries (PEDL), a Foreign, Commonwealth & Development Office (FCDO) funded programme run by the Centre for Economic Policy Research (CEPR).

Role of Firms in Wage Dispersion : Evidence from a developing country*

Yaya Diallo[1,3], Ibrahima Sarr[1], and Idrissa Diagne[2]

[1]DESEC

[2]ANSD

[1,3]McGill University

April 2022

This paper aims to extend our knowledge of wage dispersion to developing Countries. For this purpose, we built the first matched employer-employee database in a sub-Saharan African country (Senegal). Using the traditional two fixed effects model (a.k.a AKM), we find that firms explain 28.3% of the wage dispersion for Senegalese men sample and 28.6% for Senegalese female sample. The share of the variance explained by firms is much larger in Senegal than what the literature documented for High-Income countries Card et al. (2018).

*We thank the Agence Nationale de la Statistique et de l'Analyse Economie-ANSD (Statistic Agency of Senegal) for the collaboration in the implementation of this employer-employee database. We also thank all partners: (Revenue Agency(DGID), Retirement Agency (IPRES), Labor Statistics Division(DSTE), Security Agency Fund (CSS)) for their participation of the workshop on information system on employment. Special thanks to Insa SADIO (National School of Statistics-ENSAE) and Catherine MONOUE KONGA for their help during the workshop. This document is an output from the research initiave « Private Enterprise Development » (PEDL), a programme funded jointly by the Centre for Economic Policy Research (CEPR) and the Department for International Development (DFID), contract reference PEDL_LOA_8284_Sarr. The views expressed are not necessarily those of CEPR and DFID.

JEL codes: J31, J62, C23

Keys words: Firm wage premium, Sorting, Wage dispersion, Matched employer-employee data.

1 Introduction

What part of the wage inequality is explained by firms? In the standard competitive labor market model, firms take market wages as given, and firm-specific heterogeneity only influences the hired worker, but not the level of pay of any particular worker. A number of authors focusing on wage inequality have recognized and empirically proved this perspective of the standard competitive labor market model (Katz, 1986; Katz et al., 1989; Krueger and Summers, 1988). This view of the standard competitive labor market model stands in stark contrast to literature the Industrial Organization literature, which typically models markets as imperfectly competitive (Tirole, 1988; Pakes, 2017) and the fact that wages differ substantially among equally skilled workers working in different places.

Since Abowd, Kramarz, and Margolis 1999 (a.k.a AKM model) paper, which proposed a statistical model to quantify the contributions of workers and firms to earnings inequality, a growing literature using this model has emerged and widely used across countries (Card et al., 2013, 2016, 2018; Macis and Schivardi, 2016; Lavetti and Schmutte, 2016). [1] These previous studies have emphasized the importance of the workplace component in wage inequality and the existence of firm-specific pay premiums. These studies indicate that firm's effects explain 15 to 20% of overall wage variation (see Card et al. (2018) for the summary). Such empirical findings rely on matched employer-employee data from High Income-Countries (Card et al., 2018, p. 23). To estimate both firm and worker effects, researchers need dynamic matched employer-employee data. The increasing availability of such data in Western Countries has provided opportunities to explore several issues in labor economics, including the role of firms in wage dispersion. However, such findings are rare in Low-Income Countries (LICs) due to the lack of data linking employees to employers. To address this issue and provide first evidence in LICs, we build a new matched Employer-Employee Database in Senegal (Senegalese Employer-Employee Database - SEED).

[1] Some authors have used the AKM approach in different contexts: teachers and students(Rockoff (2004)), hospitals and patients(Finkelstein et al. (2016)).

In this paper, we make use of a newly matched employer-employee database to investigate the role of firm effects in the variance of earnings of formal sector workers in the formal sector in one of the developing countries. As underlined above, several recent publications in High Income Countries found that the firm-specific pay premium plays a key role in wage differential (Card et al., 2018). As the structure of labor markets between low- and high-income countries could differ in many ways due wage settings and market frictions, this paper provides a substantive contribution to this literature by extending our knowledge on the role of the firm characteristics on earning in LICs. As many African Countries have similar labor market structures, this new linked employer employee presents a test case for Sub-Saharan Countries about the role of firm-specific heterogeneity on wage inequality as well as other labor market issues.

The first contribution of this study is the construction of a new matched employer-employee database in a Low Income Country. To do so, we linked several administrative data sources from different Senegalese governmental agencies. Firstly, the employer database comes from the National Statistics Office (ANSD) and the Revenue Agency (DGID). Each firm has an unique identifier, which is National Identification Number of Companies and Associations (NINEA).[2] Secondly, the employee database comes from multiples sources: Retirement agencies (IPRESS - Institut de Prévoyance Retraite du Senegal), Social Security Fund (CSS-Caisse de Securité Sociale), and Revenue Agency (DGID-Direction Générale des Impots et Domaines). The employee data contains information on demographic characteristics (age, gender, marital status, etc.), employment and income information such as salaries, as well as information on children (age, gender). Finally, we linked these two administrative data through firm identifiers (NINEA). The data covers the period 2008-2020. To thee best of our knowledge, the SEED will be the first dynamic linked employer-employee database in a Sub-Saharan African Countries. Secondly, this paper will shed light on the importance of work place component on earning

[2]The national identification number of companies and associations (NINEA) is the unique number by which a firm or any organization is registered in the directory of companies, establishments and associations. The NINEA is an essential identifier to facilitate its administrative procedures. Any formal firm in Senegal has a NINEA.

in a Sub-Sahara African country.

The paper excludes the agricultural and makes use only the period 2015-2020. The analysis also only considers firms having at least 5 nonsingleton observations and operating for more than more years. The AKM's model of additive fixed effects for firms and workers (Abowd, Kramarz, and Margolis, 1999) is estimated using our new linked Employer-Employee for both male and female. The results indicate the important role of firms in wage inequality and earnings. The firm-specific wage premium account for 28.3% of the Senegalese men's wage variance and 28.6% of Senegalese female's wage variance. However, we observe a low worker mobility (4% per year) compared to High Income Countries (around 26% in Canada).

The remaining paper is organized as follows. Section 2 describes the data and the role of the formal sector in Senegal. Section 3 assesses the role of firms/establishments in earnings using two methods of the variance. The first method is "Plug-in Estimator", which is biased but used by economists during the two last decades to estimate variances components.[3] The second method is new approach proposed by Kline, Saggio, and Sølvsten (2020), a.k.a KSS correction. Section 4 presents and discusses the results.

[3]Traditional variance component estimators are predicated on the assumption that the errors in a linear model are identically distributed draws from a normal distribution.

2 Country background and Employment Context

During the last years, Senegal has experienced remarkable economic growth. The economy grew by more than 6% per year between 2008 and 2018 (ANSD). The dynamism of the informal sector in creating value-added and employment is powerful across African Countries.

In Senegal, as elsewhere, employment strengthens during periods of economic expansion and vice versa. From 2000-2017, the share of employment elasticity to growth was 0.55 (Mbaye et al., 2021), suggesting that a one percentage point increase in economic growth is associated with a 0.55 percentage point increase in employment. The Senegalese employment elasticity of 0.55 is higher than the average in African countries (0.41) and very close to the ideal of 0.7 (Mbaye et al., 2019).

Moreover, more than a third (43.3%) of the working-age Senegalese population had a job in the fourth quarter of 2020 (ANSD, 2021). The employment rate is higher in urban areas than in rural areas (47.7% vs. 37.9%) (ANSD, 2021). More considerable differences are observed between men and women. Indeed, the employment rate of men is 56.2% compared to 31.1% for women(ANSD, 2021).

In Senegal, the informal sector is an essential source of employment, accounting for 49% of jobs and contributing about 42% of Senegal's GDP in the non-agriculture sector (ANSD, ENSIS - 2011)[4]. In agriculture and forestry, the share of the informal sector in value-added is close to 100% (Benjamin et al., 2012), justifying the exclusion of this sector in our analysis, as we mentioned later.

Like other Sub-Saharan Countries, firms' owners in the Senegalese informal sector are mainly men (ANSD, 2011). Indeed, only 20% of owners in the informal sector are female (ANSD, 2011). The average age of entrepreneurs is 40-year-old, and owners with a low level of education dominate the informal sector because only 22.5% have reached the secondary level. The owners' lack of education could explain why they do not produce a formal accounting and do not register with the Senegalese revenue agency or the social security agency. Since 2010, the government of Senegal has introduced several reforms to

[4]Informal survey: Enquete Nationale sur le Secteur Informel au Sénégal (ENSIS-2011)

encourage informal businesses to move to the formal sector by registering in administrative files. The first crucial step in formalizing informal firms is to have a unique identifier Number called NINEA (Numéro d'Identification National des Enterprises et Associations) provided by the National Identification Center (CNI-Centre d'Identification National). Despite these governmental policies, only 8% of informal firms have a NINEA and 2% are registered in the other administrative agencies (Retirement Agency (IPRES), Social Security Agency (CSS)). Beyond registration in administrative files, informality is characterized by failing to produce financial accounts according to SYSCOA standards (West African Accounting System).

There is a strong relationship between the formal and informal sectors. For example, much of the inputs in manufacturing come from the agriculture sector, where almost 100% of firms are informal (Benjamin et al., 2012). As we are using only firms in the formal sector to construct the matched employer-employee, we must underline the interdependence between these two sectors and be careful in comparing our findings with existing studies using data from developed countries.

3 Empirical strategy

3.1 Identification and Estimation

3.1.1 AKM Models

The concerns of wage inequalities have received sustained attention among economists and policy makers in both developed and developing countries during last decades. In developed countries, the main interest is to understand the source of the inequalities of earnings in order to implement more appropriate policies to reduce the poverty.

Since the availability of matched employer employee dataset, there is a growing literature on the role of firm in wage inequality in High Income Countries (Card et al., 2018). In this section, we used detailed matched employer-employee administrative data from Senegal to quantify the contribution of firms-specific wage premiums on wage inequality in Senegal. We used the traditional AKM model to find the firm and worker fixed effects (Abowd, Kramarz, and Margolis 1999). The log-earnings can be expressed as a sum of worker effects, firm effects, covariates, and idiosyncratic error terms. The estimates will be used to decompose the variancce of log-earnings and quantify the contribution of each component: worker, firm and sorting of high-wage workers. Due to lack of data in low-income countries, our results will provide the first contribution of firm-specific wage premiums to wage inequality in a developing country. Over the last decades, the AKM model has been widely used in developed countries as summarized by Card et al. 2018.

3.1.2 Traditional AKM Framework

Since Abowd et al. 1999, several studies have discussed the identification problems of two ways fixed-effect model. The goal is to disentangle the components of wages variation attributable to workplace-specific and worker-specific heterogeneity (Abowd et al., 2002). In the matched employer-employee data, we have N_t person-years observations on N workers and J firms. We assume that the data generation of log monthly real wage $\ln y_{it}$ of individual i at time t is specified as follows:

$$\ln y_{it} = \alpha_i + \psi_{J(i,t)} + X_{it}\beta + r_{it} \tag{1}$$

where the worker component is α_i, the workplace or firm component is $\psi_{J(i,t)}$, and the index of time-varying observable characteristics is $X_{it}\beta$. The error term r_{it} , which capture unobserved factor, is a combination of the separate random effects or person-specific job match component $\lambda_{iJ(i,t)}$, unit root component ξ_{it}, and a transitory error ϵ_{it} as indicated in the following equation:

$$r_{it} = \lambda_{iJ(i,t)} + \xi_{it} + x_{it}\beta + \epsilon_{it} \tag{2}$$

Following Abowd et al. (1999) and Card et al. (2013), the worker component α_i in equation (1) can be interpreted as a combination of skills and other factors that are rewarded equally across firms. The firm-fixed effect, ψ_J, is interpreted as the proportional pay premium that is paid by firm J to all employees. Some studies suggest that this firm-fixed effect represents the rent-sharing - an efficiency wage premium, or strategic wage posting behavior (Burdett and Mortensen, 1998). The variable in X_{it} are year dummies, age, education, and some interactions. As suggested by Card et al. (2018), we normalize the age (age-40) and allow for age to be nonlinear by including quadratic and cubic terms. The index $X_{it}\beta$ is interpreted as a combination of life cycle and aggregate factors that affect worker productivity at all jobs. As well know in the literature of AKM model, estimating equation (1) by OLS can yield to bias coefficients. The main concern comes from the assumption of exogenous mobility of workers for identification of firm effects and worker effects. We assume the following condition hold:

$$\mathbb{E}\left\{r_{it}|\alpha_1, ..., \alpha_N, \psi_1, ..., \psi_N, X_{11}, ..., X_{NT}\right\} = 0 \tag{3}$$

Assuming that (3) holds in equation (1) means that the residuals r_{it} is independent of past and future firm indicators. Thus, there is no problem of endogenous mobility of workers and state dependence. The second concern is the calculation of the contributions of firm effects and sorting in the variance of earnings.[5]

[5]AKM uses the OLS estimates to decompose the variance without any correction. This method is know as "plug-in" estimate of the variance.

3.1.3 Variance Decomposition

By estimating the additive model in (1) under condition (3) holds, the estimates from the AKM model will used to quantify the role of firm, worker and sorting in setting wages. The variance of log wage is defined as follows [6]:

$$\begin{aligned} Var(y_{it}) = & \underbrace{var(\alpha_i)}_{\text{Person effects}} + \underbrace{var(\psi_{J(i,t)})}_{\text{Firm effects}} + \underbrace{var(x_{it}\beta)}_{\text{Covariates}} + 2\underbrace{cov\left(\alpha_{it}, \psi_{J(i,t)}\right)}_{\text{Sorting}} \\ & + 2cov\left(\psi_{J(i,t)}, x_{it}\beta\right) + 2cov\left(\alpha_{it}, x_{it}\beta\right) + cov(r_{it}) \end{aligned} \tag{4}$$

The AKM's decomposition provides both the firm specific wage premiums $var(\psi_{J(i,t)})$ and a term reflecting the covariance of the worker and firm effects, $cov\left(\alpha_{it}, \psi_{J(i,t)}\right)$. The last term indicates the sorting of workers to firms in the overall wage inequality. For example, if the term $cov\left(\alpha_{it}, \psi_{J(i,t)}\right)$ is positive, this means that workers with higher earnings ability are more likely to work at higher-paying firms. The contribution of the firm component to total wage variation as illustrated in equation (4), $cov(\psi_{J(i,t)}, \ln y_{it})$, contains three components: the variance of firm effect ($var(\psi_{J(i,t)})$), the sorting component ($cov(\psi_{J(i,t)}, \alpha_{it})$), and the covariance term $cov(\Psi_{J_{i,t}}, x_{it}\beta)$.

3.1.4 Limited mobility bias

After the estimation of model 1, the second concern comes from estimating the contributions of firms effects and sorting components in the decomposition (4). The estimated firm effect, $\hat{\psi}_J$, from equation 1 is unbiased but has some noisy to the additional true firm effect, ψ_J.

$$\mathbb{E}(\hat{\psi}_j) = \mathbb{E}(\psi_j - \hat{\tau}_j) = \psi_J$$

As point out by Lachowska et al. (2022), the estimation error in $\hat{\psi}_j$ will lead to bias if we are interested by the variance. To fix idea and for simplicity, let assume that $\bar{\psi}_j = 0$.

$$\mathbb{E}(\hat{\psi}_j^2) = \mathbb{E}\left[(\hat{\psi}_j - \psi_j + \psi_j)^2\right] = \psi^2 + \underbrace{var(\hat{\psi}_j)}_{\textbf{bias}} \tag{5}$$

[6] Also, can be expressed as follow :

$$Var(y_{it} - x_{it}\beta) = \underbrace{var(\alpha_i)}_{\text{Person effects}} + \underbrace{var(\psi_{J(i,t)})}_{\text{Firm effects}} + 2\underbrace{cov\left(\alpha_{it}, \psi_{J(i,t)}\right)}_{\text{Sorting}}$$

Where $var(\hat{\psi}_j)$ represents the squared standard error of the estimated firm effects, $\hat{\psi}_j$. In this case, the expected estimator $\mathbb{E}(\hat{\psi}_j^2)$ differs from the true variance component ψ^2 because of the bias introduced by $var(\hat{\psi}_j)$. The bias in equation (5) is also know as "Limited mobility bias". Andrews et al. 2008 was the first to explain that the bias arise from insufficient number of movers in the firm. The presence of the "limited mobility bias" in the decomposition of the variance tend overstated the variance of firm effects while the covariance between worker and firm tend to be understated. A recent paper of Lachowska, Mas, Saggio, and Woodbury 2022 discussed some empirical methods to correct "limited mobility bias". However, there is no yet consensus how to detect and correct this bias. The following section will presents the data and investigates the mobility of workers in our matched employer-employee. To correct the bias, we will use the KSS correction (Leave-Out Correction).

3.2 Data and measurement

The estimation of the AKM's model present in 3.1 requires an employee-employer database. We make use of several Senegalese administrative data sources to construct the first employer-employee in Africa. The database covers the period 2008 to 2020. The construction of this matched database comes as a results of combinations of different datasets obtained from three Senegalese government bodies: National Statistics Agency (Agence Nationale de la Statistique et de la Démographie - ANSD), Pension Fund agency (Institut de Prévoyance Rétraite du Sénégal - IPRES), Social Security Fund (Caisse de Securité Sociale - CSS).

First, the information on firms mostly comes from Pension Fund agency, IPRES. We complement information on firms with the ANSD's Economic and financial database (BDEF-Banque des Données Economiques et Financieres). This merging is done through unique identifiers. Any enterprise/organization operating in Senegal has a unique ID called NINEA (National Identification Number of firms and Associations). The NINEA has to be used in all administrative procedures. The merged employer database contains information on firms' annual activities such as value-added, employment, payroll, profits, and other characteristics (Industry, location, date of creation).

Secondly, constructing the employee's dataset mainly relies on the information from the Pension Fund agency. This data provides information on an individual's employment (annual earnings, job start date, job termination date). The dataset also contains workers' demographic characteristics such as occupation, age, and gender. This data is complemented with information from other government agencies including the Pension Fund Agency (IPRES - Institut de Prévoyance Rétraite du Senegal), the Social Security Fund (CSS- Caisse de Securité Sociale), and the Revenue Agency (DGID- Direction Générale des Impôts et Domaines). The datasets are linked through unique person identifiers.[7].

Thirdly, the Record of employment (ROE) provides information on the employment histories of each worker, including the type of the contract and the reason for permanent separation between employer and employees. This data comes from the Labor Statistics

[7]We identified workers using the national identity card number

Labor Division (Direction des Statistiques du Travail-DSTE).

For this paper, we restrict our sample to all workers aged 20-60 from 2015 to 2021. We exclude the agriculture sector as many jobs are informal and poorly informed. Indeed, in agriculture and forestry, the share of the informal sector in value-added is close to 100% (Benjamin et al., 2012), justifying the exclusion of this sector in our analysis. Our analysis only focuses on individuals aged between 18 and 60 with an annual salary of at least CFA 100,000 (US 200). This database does not have information on individuals who leave the formal sector. In this case, there is no possibility to distinguish moves from the informal sector or non-employment. If the individuals hold multiple jobs during the year, we consider the primary job which is the firm that paid the highest earnings for that year. Also, in Senegal, most workers are paid monthly, and most jobs are full-time, i.e. 35 hours per week. In our data, there is no possibility to distinguish full-time from partial time.

Table 1 provides a summary of our sample's main variables of interest by gender over the period 2015-2020. The full database has 712,184 person-years. This corresponds to 233,523 workers spread over 6,954 firms during the period 2015-2020 (column 1, Panel A of Table 1). The proportion of females is 28% of the total sample.

We restrict each sample to its most significant connected subset of firms in panel B for identification purposes. We used the worker's flow to find the connection between firms. The subset of connected firms is identified by the mobility of workers (Figure 1). For example, A and B are connected if a worker moves from firm A to B. Moreover, if a worker moves from B to C, firm A is connected to firm C through B. In this case, firms A, B and C are in the same connected subset. When we restrict the entire sample to the largest connected subset of firms, the number of total firms decreases from 6,954 to 1,593 (i.e. a decrease of 77%), while the total number of workers decreases by 23% – from 233,523 to 179,770. Our analysis will only use the largest connected subset of firms to estimate person and firm effects. The largest male-connected subset contains 1,275 firms. We called male-connected firms because we used only male mobility between firms to identify the largest connected firms. For the female sample, the largest connected subset contains 413 firms. In the largest connected subset of firms of the female sample, female workers represent 29.1% of employees.

Table 1: Summary Statistics of all samples

Variable	**All**	**Female Workers**	**Male Workers**
Variable	**(1)**	**(2)**	**(3)**
Panel A: Full sample			
# workers-years	712,184	187,810	524,374
# Workers	233,523	62,294	171,229
# Firms	6,954	1,979	4,924
Mean log annual earnings	14.53	14.57	14.52
Variances log annual earnings	0.80	0.79	0.80
Average experience (in years)	7.18	5.97	7.61
Panel B: Connected set of firms			
# Workers-years	541,156	96,548	385,593
# Workers	179,770	33,390	128,137
# Firms	**1,593**	**413**	**1,275**
Proportion of movers	**7.08%**	**6.12%**	**7.31%**
Share of female	.266	.291	.261
Mean log annual earnings	7.59	7.64	7.57
Variances log annual earnings	0.91	0.90	0.91
Average experience (in years)	6.32	4.78	6.75

Panel A: Firms in non-agriculture sector with more than 5 employees per year

Panel B: Firms belonging to the largest connected set of firms through job to job mobility. We calculate the share of female based on the connected set of firms. For example, for female sample, we found there are 413 connected firms through female mobility and the proportion of female in these firms is 29%.

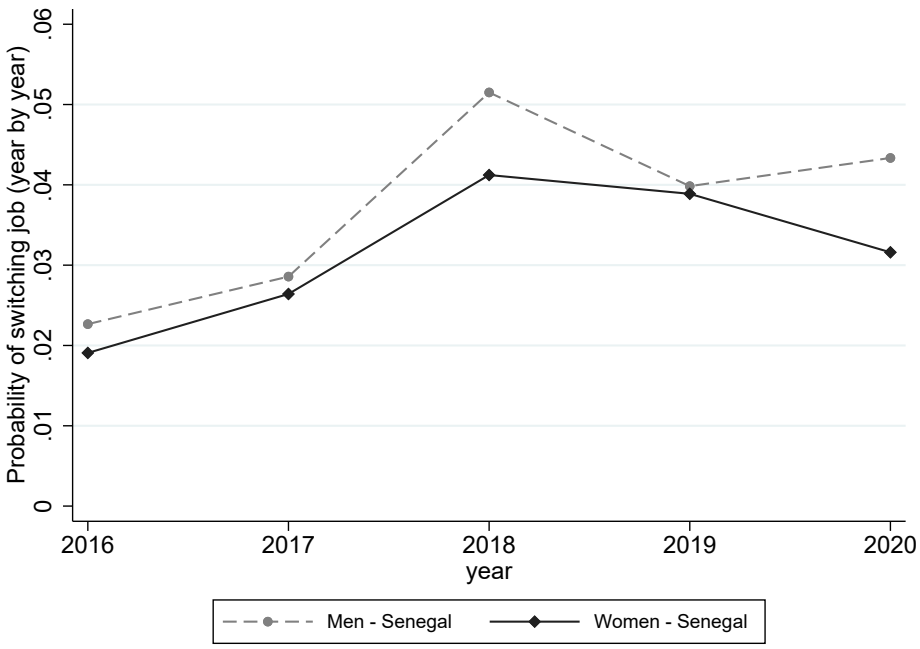

Figure 1: Job mobility of workers (female and male)

Figure 1 shows that the job mobility — the movement of employees across firms — in Senegal for the entire sample is on average 4% per year over the period 2015-2020. The figure also indicates that women are less likely to change jobs than men, and the mobility of men and women might be explained by different factors. The following section present the result of the decomposition and compare our result with findings summarized in Card et al. (2018).

4 Results

4.1 Results of variance decomposition.

The model (1) was estimated using our new matched employer-employee covering the period 2015-2020. We used the AKM decomposition as indicated in Table 2. Panel A provides the summary statistics of the samples as described in section 3.2. Panel B shows the results of the AKM decomposition while panel C shed light on the correlation between firms effects and person effects. Column (1), (2), and (3) displays the results for the full sample, the women worker sub-sample, and men worker sub-sample, respectively.

As shown in Panel A of Table 1, there are 541,156 person-years observations corresponding to 179,770 workers spread over 1,593 firms for the period 2015-2020 in the largest connected subset of firms for our full sample. The number of person-year observations for male workers (385,593 in Column 3) is much larger than for female workers (96,548 in Column 2). There are 33,390 female workers spread over 413 firms (Column 1, Table 1). In men sample, we have 128,137 workers working in 1,275 firms in the largest connected subset of firms for the men-sample. Thus, we are only able to estimate the firm effects of 413 firms for female and 1,275 firms for men.

Panel B in Table 2 quantifies the role of firms in the variance of workers' earnings using the traditional "AKM" decomposition from equation (4) and the alternative method of decomposition proposed by Kline et al. (2020) [8]. Our results are different from what we found in the literature in high-income countries (more detail in section Discussion). The estimates of the contribution of each component of the variance are presented.[9] For example, the role of firm to earnings inequality is $var(\psi_{J(i,t)})/var(y_{it})$ for the connected

[8] The sum of firm-related components is equal $Cov(\psi_{J(i,t)}, \ln y_{it}) = var(\psi_{J(i,t)}) + cov\left(\alpha_{it}, \psi_{J(i,t)}\right) + cov\left(\psi_{J(i,t)}, x_{it}\beta\right)$

[9]

$$
\begin{aligned}
1 = & \underbrace{var(\alpha_i)/var(y_{it} - x_{it}\beta)}_{\text{person effect}} + \underbrace{var(\psi_{J(i,t)})/var(y_{it} - x_{it}\beta)}_{\text{firm effect}} \\
& + 2\underbrace{\frac{cov\left(\alpha_{it}, \psi_{J(i,t)}\right)}{var(y_{it} - x_{it}\beta)}}_{\text{Sorting}} + cov(r_{it})/var(y_{it} - x_{it}\beta)
\end{aligned}
\tag{6}
$$

Table 2: Variance decomposition

	All	female workers	male workers
	(1)	(2)	(3)
Panel A: Largest connected subset			
# Firms	**1,593**	**413**	**1,275**
Variances log annual earnings	0.91	0.90	0.91
Panel A: Variance decomposition (Plug-in Estimator)			
	Share	**Share**	**Share**
var(person)	0.785	0.9	0.781
var(firm)	**0.265**	**0.286**	**0.283**
2cov(person, firm)	-0.181	-0.333	-0.192
var(residuals)	0.131	0.147	0.128
Panel B: Variance decomposition (KSS Correction)			
var(person)	0.657	0.682	0.661
var(firm)	**0.178**	**0.104**	**0.203**
2cov(person, firm)	0.022	0.052	-0.003
var(residuals)	0.143	0.162	0.139
Total firm related components (KSS)	0.188	0.131	0.201
Total person related components (KSS)	0.669	0.707	0.660
Panel C: Correlations			
Corr(Fe male, Fe female)		**0.9602**	

Note: person: Person fixed effect; **firm**: Firm fixed effect; **xb**: Covariates.

The total firm component is var(firm)+Cov(person, firm). The total person component is var(person)+Cov(person, firm).

subset of firms. For the full sample (Column 1), the role of firms is 26.5% of the variance of workers' earnings. The results of the male sample are quite similar to the full sample (Column 3) i.e firms effect account for 28.3% of the variance of earnings (drop to 20.3% with KSS correction). For female workers, 28.6% of the variance of earnings is attributable to the role of firms, which is slightly different than male workers (compared to 10.4% with correction). We found a important role of the covariance of firm and worker effects for the full sample (-22%) and male sample (-18.6%). As discussed in subsection 3.1.4, the estimates of an AKM model can be bias due to limited mobility of workers across firms, a.k.a Limited mobility bias.[10] These biases can be particularly severe when the number of workers transitioning between different employers is relatively low (Kline et al. 2020).

The KSS correction, which is preferred decomposition is presented in Section C. The correction reduced the role of firm in explaining the wage dispersion and the sorting component becomes positive suggesting high-wage workers sort into low-paying firms. We find that the total role of firms (wage premium and sorting component) account for 20.1% of the variance of wage for male sample against 13.1% for female. The last row of Panel C of 2 shows that men and women have quite highly correlated firms effects (0.96), indicating that that higher paying-firms for men are also higher paying firms for women.

4.2 Discussion.

For comparison purposes, we use the traditional decomposition method, the "Plug in estimator". During the last two decades, a growing literature documenting the additive worker and firm effects wage model proposed by Abowd, Kramarz, and Margolis 1999 (a.k.a AKM's model) using linked employer-employee data. As discussed in the Introduction, this literature finds that firm effects account for between 15 and 25% of the earnings variance in High-Income Countries. Our result suggests that the role of firms in wage setting is much more significant in Low-Income Countries compared to what the literature has documented in High-Income Countries (see Card et al. (2018) for more detail).

[10]During the last two years, some studies have investigated for bias correction Bonhomme et al. (2020); Lachowska et al. (2022); Kline et al. (2020). There is no yet consensus how to correct the Limited mobility bias

The second part of the discussion focuses on the empirical evaluation of the additivity of the AKM model has been of interest in the literature. To examine this assumption empirically, we follow Card et al. (2013). One way to test this assumption is to analyze the residual mean for different groups of workers employed in different firms. For this purpose, we divide the estimated person and firm effects into deciles, corresponding to 100 person×establishment decile cells. For each cell, we calculated the residual's mean. The mean residuals of any group of workers in any group of firms should have a mean zero if the additive assumption is correct. We calculated the cell's mean residuals for both female and male samples. We observe that the mean residuals are relatively close to 0, except for the cell of lower-skilled workers and lower-paying firms for both females and males (Figures 2 and 3). There is an exception for the following cell: 1st decile of workers × 6th decile of the distribution. Thus, there is no evidence that the high skilled workers ($10th$ decile of the distribution) earn higher premiums at the highest paying firms ($10th$ decile of the distribution). The main observation for lowest ability workers in lowest paying firms is largely negative (also found in Portuguese data by Card et al. (2018)). In Portugal, Card et al. (2018) suggests that this may reflect the impact of the minimum wage. This argument can apply to Senegal as the data used in this paper covers formal sectors where the firms have to apply the minimum wage rule. Finally, the mean cell residuals test for Germany (Card et al., 2016), Italy (Macis and Schivardi, 2016), and Portugal (Card et al., 2018) shows no evidence of violation of additive AKM-style model assumptions.

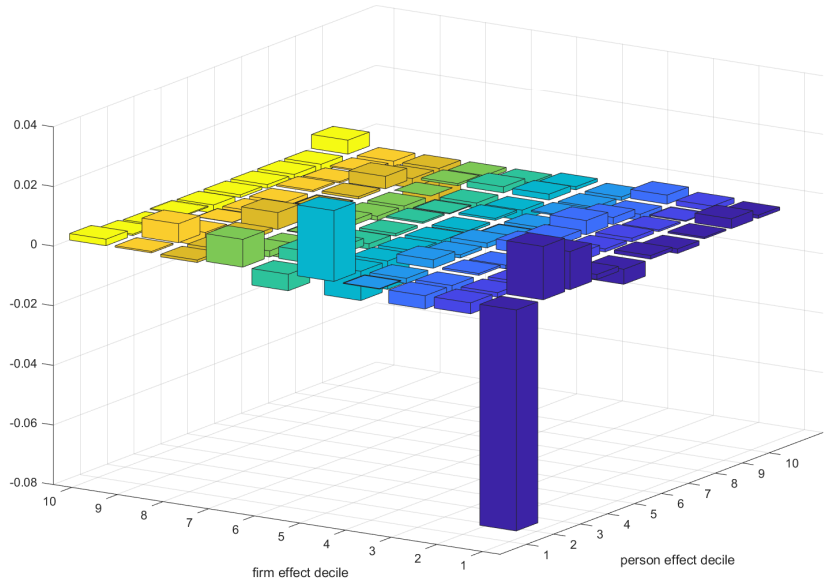

Figure 2: Mean residuals by person/firm deciles for male workers

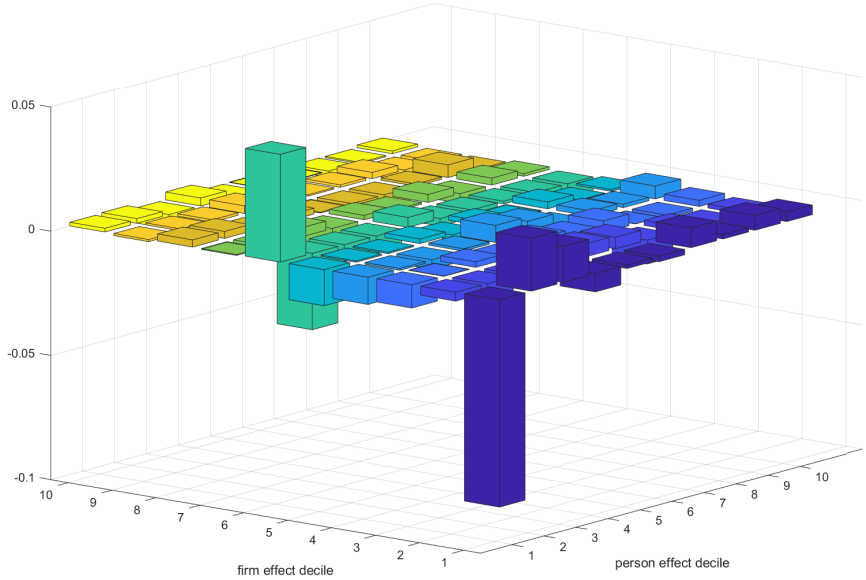

Figure 3: Mean residuals by person/firm deciles for female workers

5 Conclusion

What part of the wage inequality is explained by firms? The AKM model has been widely applied to matched employer-employee data in high-income countries to answer this question. These studies indicate that firm-specific effects premium explains 15 to 20% of overall wage variation. Such empirical results are unavailable in Low-Income Countries due to a lack of data linking firms to their employees. This study had two main objectives. Firstly, it aimed to construct a matched employer-employee database in Senegal using different datasets from separate government bodies. To our knowledge, this is the first dynamic matched employer-employee in a sub-Saharan country. Secondly, we provide the first empirical result on the role of firms in the variance of workers' earnings in one of the Low-Income countries.

Our analysis of the Senegalese data suggests that firms' role in the earnings variance for the formal sector is 28.6% for female workers and 28.6% for male workers. Our result indicates that the role of firms is much more significant in Low-Income countries compared to what the literature has documented in High-Income Countries (see Card et al. (2018) for more detail). Moreover, our results indicate that women have lower wages than men, even in companies with higher wage premiums for women and lower-paying premiums for men (figure 9).

This project produces empirical work addressing the question of wage dispersion and workers' self-selection into jobs in the context of developing countries. This empirical work is the starting point of new literature for developing countries. It helps to understand the source salary dispersion better. It shed light on the gender gap in the Senegalese economy. The finding will help to implement appropriate policies to wage equality, address the gender gap in the labour market, and promote women's labor market participation.

References

Abowd, J. M., R. H. Creecy, F. Kramarz, et al. (2002). Computing person and firm effects using linked longitudinal employer-employee data. Technical report, Center for Economic Studies, US Census Bureau.

Abowd, J. M., F. Kramarz, and D. N. Margolis (1999). High wage workers and high wage firms. *Econometrica 67*(2), 251–333.

Andrews, M. J., L. Gill, T. Schank, and R. Upward (2008). High wage workers and low wage firms: negative assortative matching or limited mobility bias? *Journal of the Royal Statistical Society: Series A (Statistics in Society) 171*(3), 673–697.

ANSD (2011). Informal survey: Enquete nationale sur le secteur informel au sénégal (ensis-2011).

ANSD (2021). Enquête nationale sur l'emploi au sénégal quatrième trimestre 2020.

Benjamin, N., A. A. Mbaye, and I. T. Diop (2012). *The informal sector in Francophone Africa: firm size, productivity, and institutions.* World Bank Publications.

Bonhomme, S., K. Holzheu, T. Lamadon, E. Manresa, M. Mogstad, and B. Setzler (2020). How much should we trust estimates of firm effects and worker sorting? Technical report, National Bureau of Economic Research.

Burdett, K. and D. T. Mortensen (1998). Wage differentials, employer size, and unemployment. *International Economic Review*, 257–273.

Card, D., A. R. Cardoso, J. Heining, and P. Kline (2018). Firms and labor market inequality: Evidence and some theory. *Journal of Labor Economics 36*(S1), S13–S70.

Card, D., A. R. Cardoso, and P. Kline (2016). Bargaining, sorting, and the gender wage gap: Quantifying the impact of firms on the relative pay of women. *The Quarterly Journal of Economics 131*(2), 633–686.

Card, D., J. Heining, and P. Kline (2013). Workplace heterogeneity and the rise of west german wage inequality. *The Quarterly journal of economics 128*(3), 967–1015.

Finkelstein, A., M. Gentzkow, and H. Williams (2016). Sources of geographic variation in health care: Evidence from patient migration. *The quarterly journal of economics 131*(4), 1681–1726.

Katz, L., L. Summers, R. E. Hall, C. L. Schultze, and R. H. Topel (1989). Industry rents: Evidence and implications. brookings papers on economic activity: Microeconomics.

Katz, L. F. (1986). Efficiency wage theories: A partial evaluation. *NBER macroeconomics annual 1*, 235–276.

Kline, P., R. Saggio, and M. Sølvsten (2020). Leave-out estimation of variance components. *Econometrica 88*(5), 1859–1898.

Krueger, A. B. and L. H. Summers (1988). Efficiency wages and the inter-industry wage structure. *Econometrica: Journal of the Econometric Society*, 259–293.

Lachowska, M., A. Mas, R. Saggio, and S. A. Woodbury (2022). Do firm effects drift? evidence from washington administrative data. *Journal of Econometrics*.

Lavetti, K. and I. M. Schmutte (2016). Estimating compensating wage differentials with endogenous job mobility.

Macis, M. and F. Schivardi (2016). Exports and wages: rent sharing, workforce composition, or returns to skills? *Journal of Labor Economics 34*(4), 945–978.

Mbaye, A. A., B. S. Coulibaly, and D. Gandhi (2019). Job creation for youth in africa.

Mbaye, A. A., F. Gueye, A. Beye, A. K. Dia, and M. Mbaye (2021). Employment creation potential, labor skills requirements, and skill gaps for young people: A senegal case study.

Pakes, A. (2017). Empirical tools and competition analysis: Past progress and current problems. *International Journal of Industrial Organization 53*, 241–266.

Rockoff, J. E. (2004). The impact of individual teachers on student achievement: Evidence from panel data. *American economic review 94*(2), 247–252.

Tirole, J. (1988). *The theory of industrial organization*. MIT press.

A Appendix A

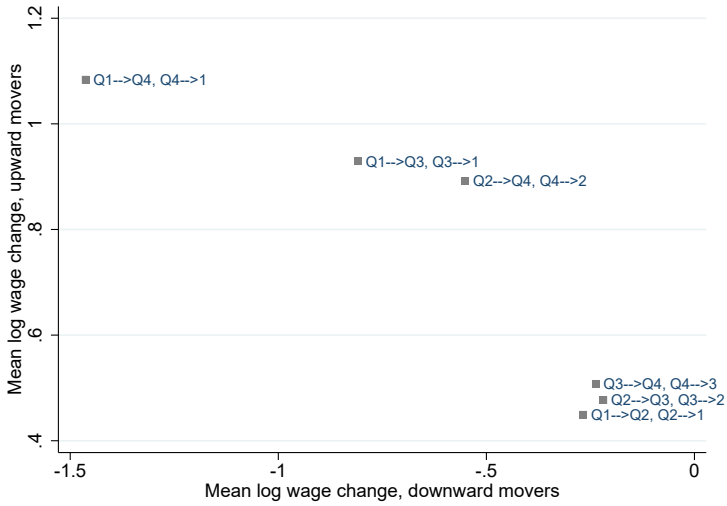

Figure 4: Wage changes for female job movers who move across firm fixed effect quartile group

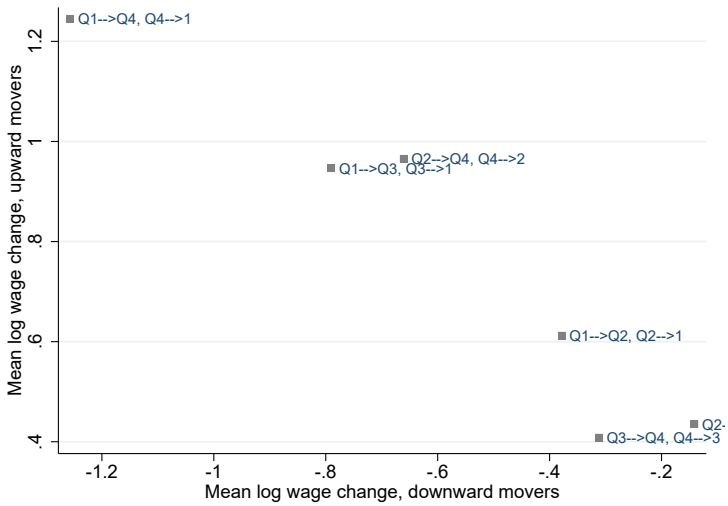

Figure 5: Wage changes for male job movers who move across firm fixed effect quartile group

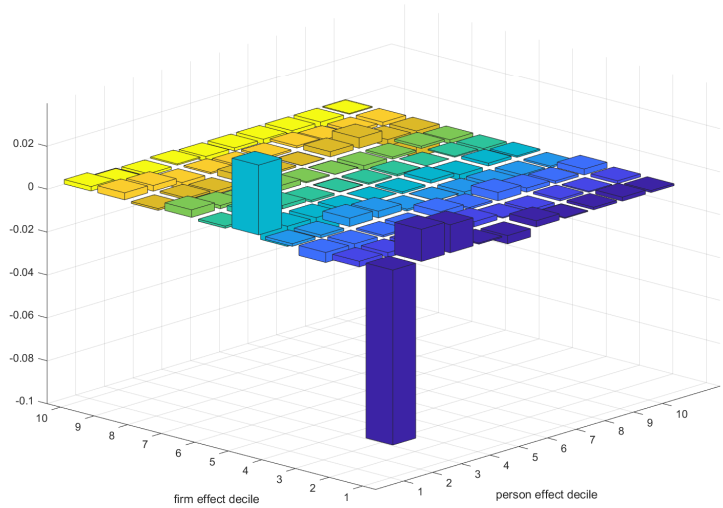

Figure 6: Mean residuals by person/firm deciles for all workers

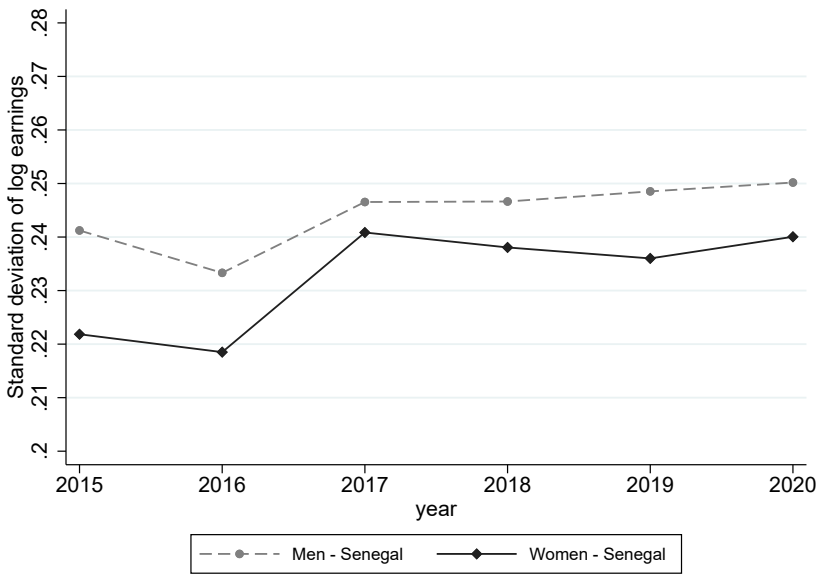

Figure 7: Trends in Wage Inequality for Male and Female Workers

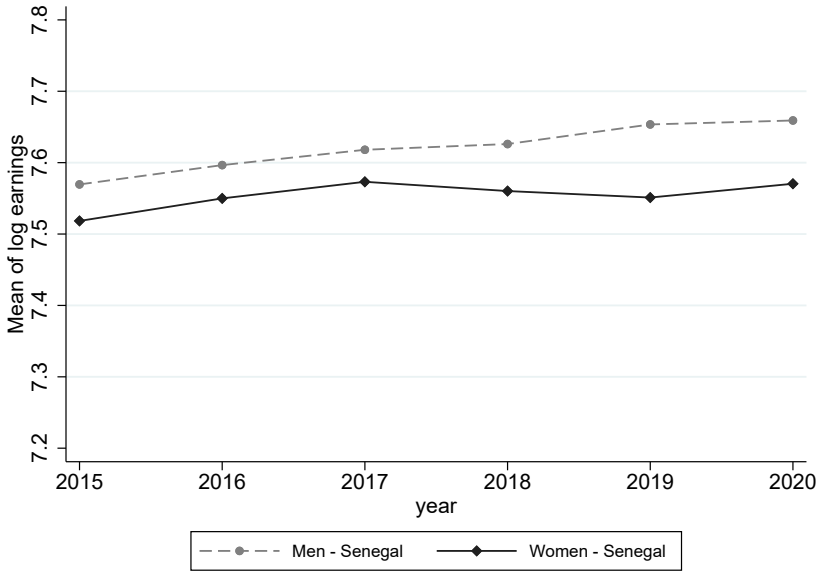

Figure 8: Trends of mean log annual earnings for Male and Female Workers

Figure 9: Trends of mean log annual earnings for Male and Female Workers

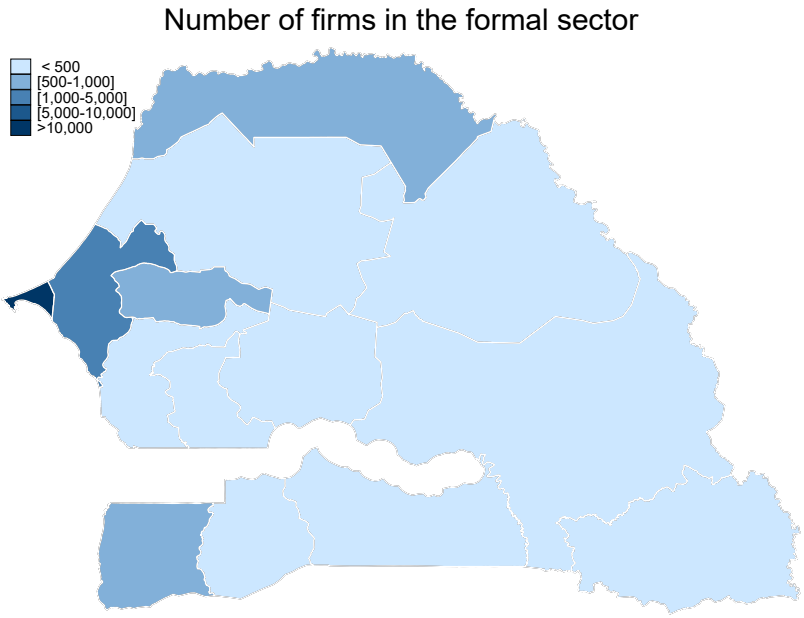

Figure 10: Number of firms by region (2020)

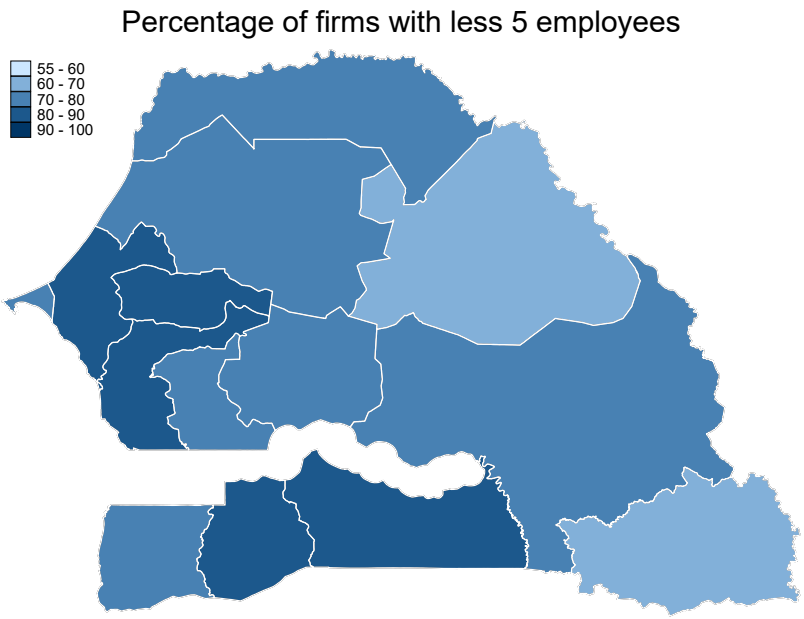

Figure 11: Percentage of firms with less than 5 employees (2020)