Understanding the Urban Earnings Growth Premium*

Guangbin Hong†

University of Toronto

March 22, 2024

Preliminary Draft

Abstract

It has been documented that larger cities foster faster earnings growth, which is an important driver for spatial earnings inequality. In this paper, I empirically investigate the sources of the urban earnings growth premium. I find that the between-firm and within-firm growth components each explain 66% and 34% of the greater returns to big city experience, respectively. Workers do not move between jobs more frequently but enjoy a steeper job ladder in larger cities. Faster within-firm learning in larger cities is mostly explained by better learning environments at the firm level. The empirical results highlight the important role of firm heterogeneity across cities in explaining the dynamic gains from working in bigger cities.

Keywords: Urban earnings growth premium, job ladder, learning, firm heterogeneity.

JEL codes: R12, R13.

*This is the second chapter of Guangbin Hong's PhD dissertation. I am deeply indebted to Joseph B. Steinberg, Nathaniel Baum-Snow, Kevin Lim, Serdar Ozkan, and Kory Kroft for their continuous guidance and support. The empirical analysis of this paper is conducted in the Statistics Canada Research Data Center (RDC). The results and views expressed in this paper are those of the author and do not represent Statistics Canada. All remaining errors are on my own.

†Department of Economics, University of Toronto. Email: g.hong@mail.utoronto.ca.

1 Introduction

Workers enjoy greater returns to work experience in big cities, which account for at least half of the urban earnings premium (Baum-Snow and Pavan, 2012; De La Roca and Puga, 2017). Therefore, uncovering the drivers of urban earnings growth premium (henceforth UEGP) is important to understand the spatial earnings structure. Previous studies have found that knowledge spillovers allow workers to accumulate human capital faster in big cities (Ellison et al., 2010; Atkin et al., 2022; Baum-Snow et al., 2023). Meanwhile, other studies have emphasized the matching benefits in big cities, helping workers move up the job ladder and earn more (Petrongolo and Pissarides, 2006; Bleakley and Lin, 2012; Eckert et al., 2022). However, little empirical evidence exists on the relative importance of the two channels in their contribution to the UEGP. Such evidence can not only help shed light on the nature of agglomeration economies but also have significant implications for aggregate outcomes.[1]

In this paper, I empirically investigate the sources of faster earnings growth in big cities. Using panel data recording the job and location history of the universe of Canadian workers, I examine earnings growth across different cities and decompose it into growth components from between-firm mobility and within-firm learning. I find that doubling the city population is associated with a 0.11 percentage point increase in annual earnings growth. Earnings growth from between-firm mobility and within-firm learning each account for 65.6% and 34.4% of the UEGP, respectively. Workers do not move between jobs more frequently in larger cities, yet the average gain from each job-to-job movement is much higher, implying a steeper job ladder. Working for larger and high-paying firms and with more skilled co-workers accounts almost entirely for faster within-job learning in larger cities.

My empirical analysis underscores the important role of spatial firm sorting in shaping the UEGP. While prior studies have emphasized the significance of matching with high-paying firms in larger cities for the urban earnings premium, I extend this understanding by showing that more productive firms in larger cities also facilitate faster earnings growth.

[1]See Duranton and Puga (2004) for the anatomy of the micro-foundations of the agglomeration spillovers, i.e. sharing, matching, and learning. And see Martellini (2019), Davis and Dingel (2019), Crews (2023) and Duranton and Puga (2023) for discussions on how local knowledge spillovers affect aggregate growth of the economy.

They not only increase the steepness of local job ladders but also cultivate faster on-the-job learning. It has been shown that heterogeneity in firms' promotion of human capital (Gregory, 2020; Ma et al., 2021) and learning benefits from co-workers (Herkenhoff et al., 2018; Jarosch et al., 2021) are crucial determinants of workers' lifetime earnings. Applying such insight into the spatial context, this paper contributes to the urban economics literature by documenting the importance of firm heterogeneity across cities in generating different local learning environments. With two-sided sorting of workers and firms across cities (Diamond, 2016; Gaubert, 2018) and greater extents of assortative matching within larger cities (Dauth et al., 2022), workers in urban areas benefit from better learning environments by working for higher-productivity firms and with higher-skilled peers.

The contributions of between-firm mobility and within-firm learning vary between workers of different skills and evolve with different stages over the life cycle. High-skilled workers benefit more from the working experience in big cities, suggesting skill-biased dynamic agglomeration benefits. More interestingly, the importance of between-firm mobility and within-firm learning varies monotonically with worker skill. Higher-skilled workers enjoy faster learning in big cities, whereas lower-skilled workers benefit more from the job ladder. In addition, the UEGP decreases over the life cycle. Workers in larger cities benefit from significantly faster learning before age 30. The steeper job ladder benefit plays a more important role in the latter stages of the life cycle. It suggests that larger cities help workers accumulate human capital faster at the start of their careers and help them find better matches to fully utilize their skills later on. Baum-Snow and Pavan (2012) document a migration pattern that college-educated workers are more likely to migrate to big cities when young and leave big cities when older. My empirical findings of skill-biased urban learning premium and the decreasing importance of human capital accumulation over the life cycle provide a rationale for such a pattern.

I also examine the earnings growth for between-city migrants. Compared to workers who do not move in the sample period, there is a stronger relationship between earnings growth and city population for between-cities migrants. Specifically, the city-size elasticity of earnings growth for migrants in their destination city is 0.46, about 3 times as large as the elasticity of 0.17 for the never-movers. The migrants not only learn faster within each firm where they are employed but also enjoy a steeper job ladder. This suggests that workers moving to bigger cities benefit more from the urban environment. One explanation for such difference is that individuals with better people skills and self-confidence tend to move to larger cities, as found in Bacolod et al. (2009) and De la Roca et al. (2023). Such skills complement the big-city environment by facilitating more social interactions and more progressive job movements.

The empirical results should not be interpreted as causal estimates of the treatment effect of the urban environment. For example, workers who self-select in larger cities may possess more specialized skills, which makes them benefit more from having access to a larger and more diverse set of firms. Nonetheless, the empirical finding of greater returns of big-city experiences for migrants provides a cautionary note for studies applying the between-city movers design to study the causal static and dynamic gains of working in large cities (e.g. De La Roca and Puga (2017) and Porcher et al. (2023)).[2] If the movers expect greater long-term returns from large city experiences, they may trade off in the short term by moving to below-average firms when they migrate for greater long-run benefits.[3] As a result, the earnings changes associated with the between-city movement could understate the static earnings gain. Moreover, movers and stayers exhibit very different earnings-experience profiles. Empirical estimates that do not distinguish between the two may mask significant underlying heterogeneity in the returns to big-city experience, potentially stemming from multidimensional worker heterogeneity, e.g. ambition, learning abilities, and social skills.

I contribute to the literature that studies the greater returns to big city experiences, pioneered by Glaeser and Mare (2001) and then studied by D'Costa and Overman (2014), Baum-Snow and Pavan (2012), De La Roca and Puga (2017) and Eckert et al. (2022). The most relevant work to this paper is Baum-Snow and Pavan (2012), who use geocoded NLSY data to compare earnings growth from job-to-job transition and on-the-job learning in large, medium, and small cities in the US. Their empirical analysis relies on a small sample and focuses on the initial 15 years of work experience. By leveraging administrative data covering the universe of Canadian workers, this paper provides a holistic view of the life

[2]Many studies in this vein use an event-study figure to test the parallel pre-trend of individuals moving to different (groups of) cities and non-migrants, following Card et al. (2013). This visualization helps illustrate the (lack of) existence and significance of time-varying earnings shocks that are correlated with spatial movements, i.e. the underlying sources of endogenous mobility.

[3]See Card et al. (2023) for the empirical evidence on this. They find that migrants tend to move to below-average firms in their destination cities.

cycle earnings growth differences across cities. The decomposition confirms the importance of both the steeper job ladder and faster human capital accumulation on the UEGP, and it sheds light on the crucial role of spatial firm heterogeneity. Although building a model to account for these facts is beyond the scope of this paper, future structural work on the spatial wage structure and its aggregate implications should take these into consideration.

The rest of the paper is organized as follows. I describe the data for the empirical exercise in Section 2. I then present three empirical facts on UEGP in Section 3, followed by discussions on the heterogeneity of UEGP across skill, age, and migration status in Section 4. Finally, I conclude in Section 5.

2 Data and Sample Construction

2.1 Data

The main data sets I use are the administrative T1 Personal Master File and T4 Statement of Remuneration Paid File. These administrative files cover the universe of tax-filing individuals in Canada, and I use the files from 2010 to 2017. The T1 file records individual-level demographic characteristics and location information. The T4 file records annual job-level information including worker and firm identifiers, annual earnings, and industry. A worker can have more than one T4 slips in a year if she works for more than one firms. For multiple job holders, I keep the job that offers the largest earnings of the year and call it the main job.[4] Nominal annual earnings are converted to 2002 Canadian dollars. I merge these two administrative files using anonymized individual tax identifiers to form an annual worker panel.

A city is defined as a Census Metropolitan Area (CMA) or a Census Agglomerate (CA) delineated in the 2016 Census of Population. The concept of CMA and CAs resembles the one of commuting zones in the U.S. – they are formed by combining a population center and adjacent municipalities with a high degree of integration with the center measured by commuting flows. To only keep those with significant sizes and sufficient labor market integration, I keep CMAs and CAs with no fewer than 15,000 full-time working individuals in 2002. I further drop one small outlier city that has average earnings greater than 150%

[4] I also compute the total annual earnings by summing up earnings from all T4 files.

of the national average. This selection process leaves me with 66 cities.

2.2 Sample construction

I restrict the *baseline sample* to full-time working individuals between the ages of 25 and 60 who live in a city. The T4 files do not include information on hours worked. Following Guvenen et al. (2021), I only include workers with annual earnings from the main job no less than the equivalent of working 20 hours per week for 13 weeks at the minimum hourly wage.[5] I exclude workers employed in agriculture (NAICS 11), mining (NAICS 21), utilities (NAICS 22), education (NAICS 61), hospitals (NAICS 62), non-profit organizations (NAICS 813), and public administrations (NAICS 92). Furthermore, the *baseline sample* only includes individuals who have earnings above the minimum threshold for two consecutive years to compute earnings growth and excludes the observations of the first year when an individual moves to a new city. This later exclusion is because this paper focuses on the reasons behind the difference in earnings growth within cities of different sizes, rather than the earnings changes caused by moving across cities.

To compare the earnings growth of between-city movers and within-city stayers, I construct a *movers sample* and a *stayers sample*. The *movers sample* includes individuals who have migrated across cities only once in the sample period. Workers who migrate more than once in a seven-year sample may have low local labor market attachment and different human capital accumulation processes. I keep the first five years since the movers are in the destination city.[6] The *stayers sample* includes individuals who have stayed in the same city throughout the sample period.

In Table A.1, I show the summary statistics of the three samples. I split the *baseline sample* into one of the largest three cities and one of the smaller cities for a first-pass comparison. I multiply the changes in log earnings by 100 to express them as percentage point changes. We can have several immediate observations from the summary statistics table. First, on average a Canadian worker experiences 1.05% earnings growth annually. The average earnings growth in large cities (1.42%) is more than twice the growth in smaller ones

[5]Canadian provinces set their own minimum wage standards. I take the lowest of these provincial standards as the national minimum wage.

[6]I do not include the observations of the movement years. The earnings changes in these movement years reflect the static benefits of migration, which are out of the scope of this paper.

(0.68%), showing the UEGP. Second, the standard deviation of annual earnings growth is 47.85% for all Canadian workers.[7] Individuals in larger cities also exhibit a slightly larger standard deviation of earnings growth, suggesting greater earnings risks. The two first observations are similar to the cross-sectional spatial earnings structure in that larger cities have both higher average earnings and greater earnings dispersions (e.g. Baum-Snow and Pavan (2013) and Dauth et al. (2022)). Third, a larger share of workers in big cities work for large firms and in high-wage industries, such as the tradeable services sector, than in small cities. This suggests that the heterogeneity in firm composition and industrial structure may play crucial roles in explaining the spatial earnings growth difference. Fourth, between-city migrants tend to be younger and experience faster earnings growth than within-city stayers. It is well-known that earnings growth is faster when workers are young. Later, I will show that this migrants-stayers difference in earnings growth is robust to controlling for the life-cycle earnings growth profile.

3 Empirical Facts

In this section, I present three empirical facts on earnings growth across cities. First, I compare average earnings growth between larger and smaller cities, establishing the UEGP. Second, I decompose city-level average earnings growth into between-firm and within-firm growth components and compare these two components across cities with different populations. Third, I assess the extent to which spatial variations in firm composition and industrial structure can explain the differences in within-firm earnings growth across cities.

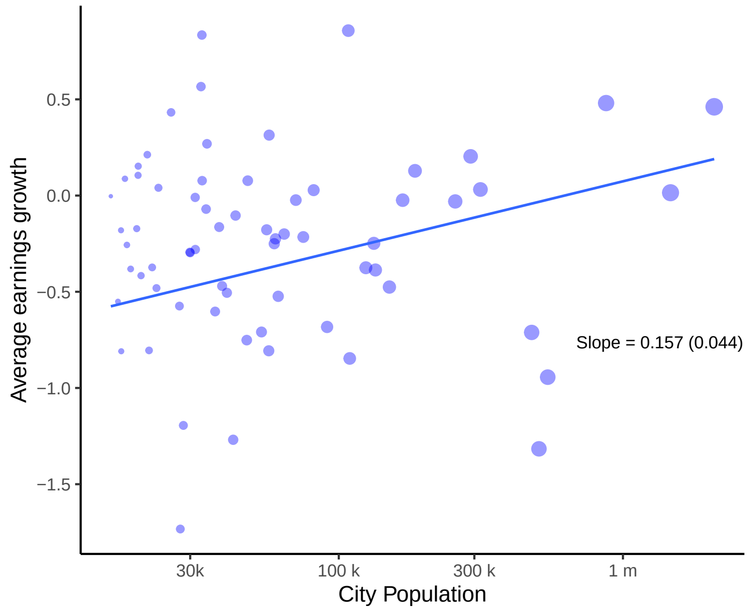
3.1 Urban earnings growth premium

I begin by associating city-level average earnings growth, denoted as $\mathbb{E}_c[\Delta w_{ict}]$, with city size, which is shown in Figure 1. The literature has documented a robust empirical pattern of a hump-shaped age-earnings profile.[8] This pattern entails low earnings for young workers upon entering the labor market, followed by an increase with age, leveling off at a certain point, and eventually declining after reaching peak earnings age. To address

[7] This is close to the standard deviation of 51% as reported in Guvenen et al. (2021) the Social Security Administration data.

[8] See Sanders and Taber (2012) for a review on literature studying the life-cycle wage growth.

Figure 1: The relationship between earnings growth and city population



Notes: Individual earnings growth used to compute the city-level growth terms have been residualized of a third-order age polynomial and year dummies. The point estimate with the robust standard error regression, which is weighted by the number of worker-year observations of each city, is reported in the figure.

the age effect, I residualize individual earnings changes by controlling for a third-order age polynomial. I follow Card et al. (2013) to assume that the lifetime earnings profile is flat at age 40. In addition, I include a set of year dummies to account for yearly macroeconomic shocks. I construct $\mathbb{E}_c[\Delta w_{ict}]$ by taking the average of the residualized individual earnings growth for each city and multiplying it by 100 to represent percentage point changes.

Figure 1 reveals a significant positive relationship between earnings growth and city population, with an earnings-growth-to-population elasticity estimated as 0.157.[9] This translates to an additional 1.1 percent increase in earnings after 10 years of experience working in a city twice as large, which is similar to the estimates of extra earnings gains from working in big cities reported in De La Roca and Puga (2017) and Porcher et al. (2023) using Spanish administrative data.[10] As reported in Hong (2024), doubling the city population is associated with a 1.6% increase in annual earnings in Canada. This cross-sectional city-size

[9]Figure 1 uses workers' earnings growth from the main job. I also examine the relationship using workers' total earnings growth and find a similar result.

[10]For example, Porcher et al. (2023) find that workers benefit from an additional 1.23 percent in earnings after 9.4 years of working experience in a city twice as large.

earnings premium corresponds to 21 years of accumulated earnings gap at an annual growth rate difference of 0.109 percentage point, if the workers start their careers with identical earnings.

Two points are noteworthy for the result. First, as mentioned earlier, only workers who have stayed in the same city for two consecutive years are included in the sample. Earnings changes associated with migration are informative of the static gains of moving to a city, which is beyond the scope of this paper. De La Roca and Puga (2017) find the initial earnings premium of big cities to be about 40% of the medium-term premium associated with 7.7 years of local experience. Second, it does not speak to differences in unemployment risks across cities, as workers who report annual earnings below a minimum threshold are excluded from the sample. By examining the earnings dynamics of U.S. workers, Guvenen et al. (2021) find that job loss risks constitute an important source of earnings risks, especially for low-income workers. In Figure A.1, I plot the unemployment rate against the city population. The figure reveals considerable geographical variation in local unemployment rates, yet no systematic relationship between unemployment rates and city population.[11] This shows that larger cities do not necessarily mitigate or exacerbate unemployment risks.

I also compare the urban earnings growth premium by sector in Table A.2. The results reveal a moderate extent of heterogeneity across sectors. On the one hand, workers in the transportation sector benefit the most from working in bigger cities, with a growth-population elasticity estimated as 0.239. On the other hand, the returns from big city experience for the workers in the retail and wholesale trade sector is only about one-third compared to transportation workers.

3.2 Decomposing the urban earnings growth premium

The greater earnings growth in large cities may be because workers take advantage of a steeper job ladder and because workers learn faster within each job. To understand how job mobility and on-the-job learning contribute to spatial disparities in earnings growth, I

[11]See Kline and Moretti (2013) and Bilal (2023) for more detailed discussions on the spatial variation in unemployment rates.

Table 1: Regression of earnings growth components on city population

Components:	Total	Between-firm			Within-firm		
		Total	m_c	$\mathbb{E}_c^m[\Delta w_{ict}]$	Total	s_c	$\mathbb{E}_c^s[\Delta w_{ict}]$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log Pop.	0.157***	0.103***	0.001	0.713***	0.054*	−0.001	0.062
	(0.044)	(0.014)	(0.001)	(0.110)	(0.032)	(0.001)	(0.038)
Constant	−2.089***	−0.973***	0.136***	−6.667***	−1.115***	0.864***	−1.308**
	(0.580)	(0.187)	(0.019)	(1.456)	(0.416)	(0.019)	(0.507)
Observations	66	66	66	66	66	66	66
R^2	0.165	0.449	0.003	0.394	0.043	0.003	0.040

Notes: The dependent variables are constructed according to equation (1). Individual earnings growth used to compute the city-level growth terms have been residualized of a third-order age polynomial and year dummies. All regressions are weighted by the number of worker-year observations in each city. Robust standard error reported. Significance level: *p<0.1; **p<0.05; ***p<0.01.

decompose city-level average earnings growth below:

$$\underbrace{\mathbb{E}_c[\Delta w_{ict}]}_{\text{total}} = \underbrace{m_c \times \mathbb{E}_c^m[\Delta w_{ict}]}_{\text{between-firm}} + \underbrace{s_c \times \mathbb{E}_c^s[\Delta w_{ict}]}_{\text{within-firm}} \tag{1}$$

where i indexes a worker, c indexes a city, and t indexes a year, $\mathbb{E}_c[\Delta w_{ict}]$ is city c's average earnings growth, $\mathbb{E}_c^m[\Delta w_{ict}]$ and $\mathbb{E}_c^s[\Delta w_{ict}]$ are city c's average earnings growth for between-firm movers and within-firm stayers, m_c is the between-firm mobility rate in city c, and $s_c := 1 - m_c$ is the staying rate in city c. Denote $j(i,t)$ as the firm j worker i is employed in year t, I define

$$\mathbb{E}_c^m[\Delta w_{ict}] = \mathbb{E}_c[w_{ic,t} - w_{ic,t+1} | j(i,t+1) \neq j(i,t), c(i,t+1) = c(i,t)]$$
$$\mathbb{E}_c^s[\Delta w_{ict}] = \mathbb{E}_c[w_{ic,t} - w_{ic,t+1} | j(i,t+1) = j(i,t), c(i,t+1) = c(i,t)]$$

and

$$m_c = \frac{\sum 1\{j(i,t+1) \neq j(i,t), c(i,t+1) = c(i,t)\}}{\sum 1\{c(i,t+1) = c(i,t)\}}$$
$$s_c = \frac{\sum 1\{j(i,t+1) = j(i,t), c(i,t+1) = c(i,t)\}}{\sum 1\{c(i,t+1) = c(i,t)\}}.$$

From equation (1), the total earnings growth is higher in a city if 1) holding the job mobility rate fixed, workers benefit more from each job movement and each additional year of experience staying in the same firm, and 2) the job mobility rate is higher if $\mathbb{E}_c^m[\Delta w_{ict}]>\mathbb{E}_c^s[\Delta w_{ict}]$ or lower if $\mathbb{E}_c^m[\Delta w_{ict}] < \mathbb{E}_c^s[\Delta w_{ict}]$. This analysis is akin to the one conducted in Ozkan et al. (2023), who studied the role of job ladder risks and human capital accumulation in shaping lifetime earnings inequality across U.S. workers.

In Table 1, I show the regressions of all terms in equation (1) on log city population. The results suggest the following points. First, comparing the slope coefficients of Columns (2) and (5) to the one of Column (1) indicates that the between-firm component and the within-firm component account for 65.6% and 34.4% of UEGP, respectively. This decomposition result should not be interpreted as a comparison of the levels of the between-firm and within-firm components. Rather, it speaks to the gradient of the two components with respect to city population and their contributions to the total earnings-growth to population gradient. Second, workers in larger cities do not move between jobs more or less frequently than workers in smaller cities, as shown by the small and insignificant slope estimate in Column (3). Using NLSY79 data, Martellini (2019) also finds that job-to-job mobility rates are similar in U.S. cities with different sizes. Third, the city-size gradients of between-firm and within-firm components are mostly accounted for by differences in earnings growth when moving between firms and when staying in a firm. Coefficients in Columns (4) and (7) imply that doubling city population is associated with a 0.50 percentage point higher earnings growth from each job-to-job movement and a 0.04 percentage point higher earnings growth from one additional year of experience staying in the same firm.

The positive correlation between city population and the gains from job-to-job movement yet population-invariant job mobility rates can be better understood in the context of a labor search model (e.g. Petrongolo and Pissarides (2006)). In larger cities, workers receive more frequent offers due to lower search frictions, and the offers are drawn from a better wage offer distribution. However, the wage distribution of employed workers is also better in large cities. Therefore, conditional on receiving an offer, the probability of accepting the offer is lower in larger cities, as workers are willing to wait longer until they receive an offer to make a favorable move. This can explain the greater gains from each job-to-job movement but not more frequent movements in larger cities.

3.3 Understanding the spatial variation in within-firm earnings growth

In this section, I investigate the role of firm heterogeneity in fostering faster earnings growth in larger cities. Workers could learn faster in larger cities due to the spatial sorting of firms and more productive firms being better learning environments (Gregory, 2020). They could also benefit more from knowledge spillovers in the denser urban environment (Atkin et al. (2022)), which is an important micro-foundation of the agglomeration economy (Duranton and Puga, 2004).[12]

To disentangle the two, I use a sample of within-firm stayers to regress individual earnings growth on log city populations while controlling for a set of industry dummies and firm characteristics. The set of firm-level controls includes log firm size, firm FE, and mean worker FE. The fixed effects are estimated from the empirical two-way fixed effects AKM equation Abowd et al. (1999), and I follow Bonhomme et al. (2019) to group firms into 10 clusters. Controlling mean worker FE allows examining the learning from co-workers mechanism explored in Jarosch et al. (2021). I also include interactions of firm characteristics and log city population to test the complementarity between city-level and firm-level environments. For example, workers may learn more from higher-skilled coworkers. Better urban amenities are likely to strengthen the learning benefits by encouraging more social interactions outside the workplace.

The results are shown in Table 2. First, I find that controlling industry dummies does not significantly affect the slope estimate, suggesting a minor role played by industrial structure. Second, all three firm characteristics included in Columns (3), (5), and (7) are significantly and positively correlated with earnings growth. At the same time, including these firm controls dramatically decreases the slope estimates of the city population. Controlling only log firm size decreases the city-size gradient almost to zero. This suggests a central role of firm heterogeneity in driving human capital accumulation differences across cities. Lastly, I find evidence of complementarity between firms and cities in facilitating human capital accumulation. This can be seen from the significant and positive estimates of the interaction terms of firm characteristics and city population in Columns (4), (6), and (8). I include all three firm controls in Column (9) and their interactions with the city population

[12] Another possibility is that employed workers in larger cities receive more competing outside offers which they leverage to bargain for higher wages with their current employer (Postel-Vinay and Robin, 2002). It can be viewed as a city-wide benefit not restricted to the firm by which a worker is employed.

Table 2: The role of firm for within-firm stayers' earnings growth across cities

Dep. Var.:	Earnings growth									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Log Pop	0.06	0.05	0.01	−0.08	0.04	−0.20***	0.02	0.13**	-0.02	-0.05
	(0.04)	(0.04)	(0.03)	(0.05)	(0.04)	(0.04)	(0.04)	(0.06)	(0.03)	(0.09)
Log Firm Size			0.13***	−0.19					0.09***	-0.08
			(0.02)	(0.14)					(0.02)	(0.15)
× Log Pop				0.02**						0.01
				(0.01)						(0.01)
Firm FE					2.24***	−3.43***			−3.43***	−4.62***
					(0.32)	(1.26)			(0.28)	(1.48)
× Log Pop						0.43***				0.09
						(0.10)				(0.12)
Avg. Worker FE							2.16***	−0.19	3.55***	2.24***
							(0.16)	(0.73)	(0.15)	(0.92)
× Log Pop								0.18***		0.10
								(0.06)		(0.07)
Industry FE	N	Y	Y	Y	Y	Y	Y	Y	Y	Y
Observations	27,859,100									

Notes: All regressions in this table use the sample of within-firm stayers. Firm size refers to the number of workers in the firm each year. Worker and firm FEs are estimated from a two-way FE earnings equation. I follow Bonhomme et al. (2019) and group firms into 10 clusters based on firms' empirical earnings distribution. Average worker FE refers to the average fixed effects of all workers employed by the firm each year. Robust standard error clustered at the city level is reported. Significance level: *p<0.1; **p<0.05; ***p<0.01.

in Column (10). Average worker FE is the only positive and significant in Column (10), indicating the predominant importance of learning from co-workers.

4 Heterogeneity analysis

4.1 Urban earnings growth premium for different skilled workers

It has been found that higher-skilled workers benefit more from big city experience than lower-skilled workers (e.g. De La Roca and Puga (2017) and Baum-Snow et al. (2018)). I investigate this skill-bias pattern in Table 3. The administrative T1 and T4 data sets do not contain education information. I use the worker fixed effect from the two-way fixed effect AKM equation to measure worker skill and I group them into four worker FE quartile groups. I regress the terms in equation (1) on log city population separately for each

Table 3: Urban earnings growth premium for different skilled workers

	Total	Between	Within		Total	Between	Within
Panel A: First quartile				*Panel B: Second quartile*			
Log Pop	0.103	0.208***	-0.105*	Log Pop	0.137***	0.117***	0.020
	(0.079)	(0.031)	(0.053)		(0.051)	(0.015)	(0.041)
Constant	−3.065***	−1.982***	-1.083	Constant	−1.640**	−1.074***	-0.566
	(1.039)	(0.414)	(0.697)		(0.672)	(0.192)	(0.543)
R^2	0.026	0.406	0.058	R^2	0.102	0.502	0.004
Panel C: Third quartile				*Panel D: Fourth quartile*			
Log Pop	0.134***	0.077***	0.057*	Log Pop	0.171***	0.028***	0.142***
	(0.037)	(0.009)	(0.032)		(0.036)	(0.008)	(0.032)
Constant	−1.416***	−0.748***	-0.668	Constant	−1.294***	-0.297***	−0.998***
	(0.490)	(0.124)	(0.421)		(0.484)	(0.102)	(0.420)
R^2	0.079	0.297	0.005	R^2	0.254	0.173	0.241

Notes: Workers are binned into four quartile groups based on worker FE estimates from the AKM equation. The regressions are weighted by the number of worker-year observations in each city-skill group bin. Robust standard error clustered at the city level is reported. Significance level: *p<0.1; **p<0.05; ***p<0.01.

worker group.

Table 3 confirms that high-skilled workers benefit more from big-city experiences than low-skilled workers, although the slope estimates are not statistically different between different groups. Zooming into the two earnings components, I find that the big city benefits heterogeneous-skilled workers in different ways. Higher-skilled workers benefit more from faster within-firm learning, whereas low-skilled workers benefit more from the steeper job ladder. Related to the results in Table 2, this cross-skill difference in the relative between-firm and within-firm contributions is closely related to the higher degree of worker-firm assortative matching in larger cities. Higher-skilled workers are much more likely to match with high-productivity firms in a larger city, which fosters greater within-firm learning and less gains from moving between firms. On the other hand, lower-skilled workers start from lower-rank firms in bigger cities and benefit the most from climbing up the steeper job ladder.

A limitation of the fixed effect approach is that the worker fixed effect contains both

Table 4: Urban earnings growth premium over the life-cycle

	Total	Between	Within		Total	Between	Within
Panel A: Age ≤ 30				*Panel B: Age 31–40*			
Log Pop	0.428***	0.224***	0.203***	Log Pop	0.104**	0.126***	−0.022
	(0.091)	(0.041)	(0.054)		(0.046)	(0.017)	(0.033)
Constant	−5.781***	−1.740***	-4.042***	Constant	−1.407**	−1.029***	0.378
	(1.196)	(0.535)	(0.716)		(0.610)	(0.226)	(0.440)
R^2	0.257	0.322	0.179	R^2	0.074	0.466	0.007
Panel C: Age 41–50				*Panel D: Age >50*			
Log Pop	0.076**	0.062***	0.014	Log Pop	0.188***	0.067***	0.122***
	(0.032)	(0.012)	(0.025)		(0.056)	(0.014)	(0.044)
Constant	−1.055**	−0.594***	-0.461	Constant	−2.414***	-0.915***	−1.498***
	(0.427)	(0.157)	(0.331)		(0.740)	(0.179)	(0.580)
R^2	0.079	0.297	0.005	R^2	0.152	0.271	0.106

Notes: The regressions are weighted by the number of worker-year observations in each city-age group bin. Robust standard error clustered at the city level is reported. Significance level: *p<0.1; **p<0.05; ***p<0.01.

innate ability and accumulated human capital. For example, a worker with low innate ability and more years of experience in a large city may be identified as a high-skilled worker. Hence, the fourth quartile group constructed by ranking worker FEs includes workers with lower innate abilities in large cities and excludes workers with higher innate abilities in smaller cities. If the UEGP positively correlates with workers' innate abilities, then the results in Table 3 would understate such a skill-biased effect.

4.2 Urban earnings growth premium over the life-cycle

in Table 4, I examine the UEGP in different stages during the life cycle. I find that workers benefit more from working in big cities when young.[13] This aligns with Baum-Snow and Pavan (2012)'s findings that workers tend to move into bigger cities when young and leave for smaller cities later. This migration pattern maximizes workers' lifetime real income by taking advantage of faster learning in bigger cities when younger and cheaper housing

[13]Recall that I already control for the average earnings-age profile when I construct the city average earnings growth terms.

Table 5: Urban earnings growth premium for between-city migrants and within-city stayers

Components:	Total	Between-firm			Within-firm		
		Total	m_c	$\mathbb{E}_c^m[\Delta w_{ict}]$	Total	s_c	$\mathbb{E}_c^s[\Delta w_{ict}]$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Between-city migrants							
Log Pop	0.459***	0.271***	0.004	1.210***	0.188**	−0.004	0.227**
	(0.105)	(0.043)	(0.003)	(0.197)	(0.078)	(0.003)	(0.098)
Constant	−3.894***	−2.152***	0.160***	−9.317***	−1.742*	0.840***	−2.046
	(1.336)	(0.545)	(0.040)	(2.495)	(0.985)	(0.040)	(1.248)
R^2	0.229	0.384	0.026	0.372	0.084	0.026	0.077
Panel B: Within-city stayers							
Log Pop	0.170***	0.108***	0.001	0.759***	0.061*	−0.001	0.071*
	(0.044)	(0.014)	(0.001)	(0.108)	(0.032)	(0.001)	(0.038)
Constant	−2.314***	−1.067***	0.123***	−7.391***	−1.246***	0.877***	−1.449***
	(0.575)	(0.181)	(0.018)	(1.431)	(0.419)	(0.018)	(0.508)
R^2	0.191	0.493	0.015	0.434	0.055	0.015	0.050

Notes: Panel A uses the *movers sample* and Panel B uses the stayers sample. The regressions are weighted by the number of worker-year observations in each city. Robust standard error clustered at the city level is reported. Significance level: *p<0.1; **p<0.05; ***p<0.01.

in smaller cities when older. Moreover, the city-size premium of within-firm growth diminishes over the life-cycle, albeit resurging after the age of 50. Given the importance of firms in explaining within-firm growth, this result implies that the match quality of the initial employer has persistent impacts on a worker's life-cycle earnings trajectory. The resurgence after 50 may reflect slower decreases in hours worked as workers approach retirement in larger cities.

4.3 Urban earnings growth premium for between-city migrants

Lastly, I compare the urban earnings growth premium of between-city movers and within-city stayers. As described in Section 2, I construct a *movers sample* keeping workers who move only once in the sample period and the five first years after a worker moves to a new city; I construct a *stayers sample* using workers who have stayed in the same city throughout the sample period. I regress the terms in equation (1) on log city population separately for the two samples.

A comparison between results in Table 5 suggests that between-city migrants benefit more from an additional year of big-city experience. The elasticity estimate in Column (1) for migrants is 2.7 times as high as the one for stayers. This is driven by being able to move to higher-paying firms, shown in Column (4), and learning faster within a job, shown in Column (7). The elasticity of the job-to-job mobility rate to the city population is also small and insignificant, showing that the migrants do not change jobs more frequently in big cities.

Many empirical studies have exploited between-city movers for causal estimates of the static and dynamic urban premium. The empirical finding of greater returns of big-city experiences for migrants provides a cautionary note for using the movers design. If the movers expect greater long-term returns from large city experiences, they may trade-off in the short term by moving to below-average firms when they migrate. Then, the earnings changes associated with the between-city movement could understate the static earnings gain. Moreover, movers and stayers exhibit very different earnings-experience profiles. Empirical estimates that do not distinguish between these two may mask massive underlying heterogeneity in the returns to big-city experience, potentially stemming from worker heterogeneity beyond a uni-dimensional vertical skill difference.

5 Conclusion

In this paper, I utilize panel data covering the universe of Canadian workers to investigate the source of the urban earnings growth premium. Specifically, I find the between-firm and within-city components each account for 65.6% and 34.4% of the greater returns to big-city experience, respectively. Greater job mobility benefits in big cities are due to greater average gains from every job-to-job movement, but not higher job-to-job mobility rates. Faster within-firm learning in big cities is mostly explained by spatial firm heterogeneity. Workers have greater earnings growth in high-wage, larger firms, and such firms systematically sort into larger cities. The returns to big city experience also vary by worker skill and different stages in the life cycle. Migrants enjoy greater growth benefits from moving to big cities than the stayers.

The findings that a worker's earnings growth depends on her own skill and the characteristics of her employer also echo the findings of Hong (2024). Both higher earnings and faster earnings growth in larger cities are tightly related to the systematic sorting of workers and firms across space. Therefore, it is crucial to account for worker and firm heterogeneity when studying both static and dynamic agglomeration effects.

The distinction between the two growth components is important for macroeconomic studies that investigate the aggregate implications of local knowledge spillovers and human capital accumulation. Such studies typically measure human capital growth differences across cities using city-level average earnings growth (e.g. Crews (2023) and Duranton and Puga (2023)). My findings suggest that neglecting the job ladder benefits of big cities will overstate the aggregate gains from local knowledge spillovers.

The empirical result highlighting the central role of firm heterogeneity in the spatial earnings growth differential has important policy implications. The local composition of firms not only matters for the earnings distribution of a city at any point in time but more importantly shapes the evolution of the distribution in the longer run. Therefore, policymakers should consider the long-term gains, in addition to the short-term gains from job creation and wage increases, when designing place-based incentives to attract high-productivity firms to economically distressed regions.

References

Abowd, J. M., F. Kramarz, and D. N. Margolis (1999): "High Wage Workers and High Wage Firms," *Econometrica*, 67, 251–333.

Atkin, D., M. K. Chen, and A. Popov (2022): "The Returns to Face-to-Face Interactions: Knowledge Spillovers in Silicon Valley," Working Paper 30147, National Bureau of Economic Research.

Bacolod, M., B. S. Blum, and W. C. Strange (2009): "Skills in the city," *Journal of Urban Economics*, 65, 136–153.

Baum-Snow, N., M. Freedman, and R. Pavan (2018): "Why has urban inequality increased?" *American Economic Journal: Applied Economics*, 10, 1–42.

Baum-Snow, N., N. Gendron-Carrier, and R. Pavan (2023): "Local productivity spillovers," *Working paper*.

Baum-Snow, N. and R. Pavan (2012): "Understanding the City Size Wage Gap," *Review of Economic Studies*, 79, 88–127.

——— (2013): "Inequality and City Size," *Review of Economics and Statistics*, 95, 1535–1548.

Bilal, A. (2023): "The Geography of Unemployment," *Quarterly Journal of Economics*, 138, 1507–1576.

Bleakley, H. and J. Lin (2012): "Thick-market effects and churning in the labor market: Evidence from US cities," *Journal of urban economics*, 72, 87–103.

Bonhomme, S., T. Lamadon, and E. Manresa (2019): "A Distributional Framework for Matched Employer Employee Data," *Econometrica*, 87, 699–739.

Card, D., J. Heining, and P. Kline (2013): "Workplace Heterogeneity and the Rise of West German Wage Inequality," *Quarterly Journal of Economics*, 128, 967–1015.

Card, D., J. Rothstein, and M. Yi (2023): "Location, Location, Location," Working Paper 31587, National Bureau of Economic Research.

Crews, L. G. (2023): "A dynamic spatial knowledge economy," .

Dauth, W., S. Findeisen, E. Moretti, and J. Suedekum (2022): "Matching in Cities," *Journal of the European Economic Association*, 20, 1478–1521.

Davis, D. R. and J. I. Dingel (2019): "A Spatial Knowledge Economy," *American Economic Review*, 109, 153–70.

D'Costa, S. and H. G. Overman (2014): "The urban wage growth premium: Sorting or learning?" *Regional Science and Urban Economics*, 48, 168–179.

De la Roca, J., G. I. Ottaviano, and D. Puga (2023): "City of dreams," *Journal of the European Economic Association*, 21, 690–726.

De La Roca, J. and D. Puga (2017): "Learning by Working in Big Cities," *Review of Economic Studies*, 84, 106–142.

Diamond, R. (2016): "The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980-2000," *American Economic Review*, 106, 479–524.

Duranton, G. and D. Puga (2004): "Micro-Foundations of Urban Agglomeration Economies," *Handbook of Regional and Urban Economics*, 4, 2063–2117.

——— (2023): "Urban growth and its aggregate implications," *Econometrica*, 91, 2219–2259.

Eckert, F., M. Hejlesen, and C. Walsh (2022): "The return to big-city experience: Evidence from refugees in Denmark," *Journal of Urban Economics*, 130, 103454.

Ellison, G., E. L. Glaeser, and W. R. Kerr (2010): "What causes industry agglomeration? Evidence from coagglomeration patterns," *American Economic Review*, 100, 1195–1213.

Gaubert, C. (2018): "Firm Sorting and Agglomeration," *American Economic Review*, 108, 3117–53.

Glaeser, E. L. and D. C. Mare (2001): "Cities and skills," *Journal of Labor Economics*, 19, 316–342.

Gregory, V. (2020): "Firms as learning environments: Implications for earnings dynamics and job search," *FRB St. Louis Working Paper*.

Guvenen, F., F. Karahan, S. Ozkan, and J. Song (2021): "What Do Data on Millions of US Workers Reveal about Lifecycle Earnings Dynamics?" *Econometrica*, 89, 2303–2339.

Herkenhoff, K., J. Lise, G. Menzio, and G. M. Phillips (2018): "Production and learning in teams," Tech. rep., National Bureau of Economic Research.

Hong, G. (2024): "Two-Sided Sorting of Workers and Firms: Implications for Spatial Inequality and Welfare," Tech. rep., Working Paper.

Jarosch, G., E. Oberfield, and E. Rossi-Hansberg (2021): "Learning from coworkers," *Econometrica*, 89, 647–676.

Kline, P. and E. Moretti (2013): "Place Based Policies with Unemployment," *American Economic Review*, 103, 238–43.

Ma, X., A. Nakab, D. Vidart, et al. (2021): *Human capital investment and development: The role of on-the-job training*, University of Connecticut, Department of Economics.

Martellini, P. (2019): "The City-Size Wage Premium: Origins and Aggregate Implications," Tech. rep., Working Paper.

Ozkan, S., J. Song, and F. Karahan (2023): "Anatomy of lifetime earnings inequality: Heterogeneity in job-ladder risk versus human capital," *Journal of Political Economy Macroeconomics*, 1, 506–550.

Petrongolo, B. and C. Pissarides (2006): "Scale effects in markets with search," *The Economic Journal*, 116, 21–44.

Porcher, C., H. Rubinton, and C. Santamaría (2023): "JUE insight: The Role of Establishment Size in the City-Size Earnings Premium," *Journal of Urban Economics*, 136, 103556.

Postel-Vinay, F. and J.-M. Robin (2002): "Equilibrium wage dispersion with worker and employer heterogeneity," *Econometrica*, 70, 2295–2350.

Sanders, C. and C. Taber (2012): "Life-cycle wage growth and heterogeneous human capital," *Annu. Rev. Econ.*, 4, 399–425.

Appendix A Additional Tables and Figures

Table A.1: Summary statistics

Sample:	Baseline			Mover	Stayer
	All	Big City	Small City		
Mean change log earnings (×100)	1.05	1.42	0.68	3.93	0.95
Std. change log earnings (×100)	47.59	47.85	47.31	54.30	47.30
Mean log earnings	10.47	10.48	10.46	10.56	10.47
Std. log earnings	0.82	0.83	0.82	0.84	0.82
Mean age	42.50	42.47	42.53	38.49	42.68
Std. age	9.61	9.51	9.72	8.81	9.60
Share in Firm Emp >50	47.5%	54.4%	40.6%	45.6%	47.7%
Employment share by sector:					
Construction	8.9%	6.9%	10.8%	8.5%	8.9%
Manufacturing	16.2%	15.9%	16.5%	13.6%	16.3%
Retail and wholesale	19.8%	19.7%	19.9%	18.4%	19.9%
Transportation	6.1%	6.4%	5.8%	5.8%	6.1%
Tradable services	32.0%	34.5%	29.6%	37.6%	31.8%
Admin, education, healthcare	6.6%	7.0%	6.3%	6.4%	6.6%
Entertainment and hospitality	10.4%	9.6%	11.2%	9.6%	10.4%
Number of worker-year observations	32563000	16279000	16284000	548000	31280000
Number of workers	7640000	3799000	3940000	224000	7304000

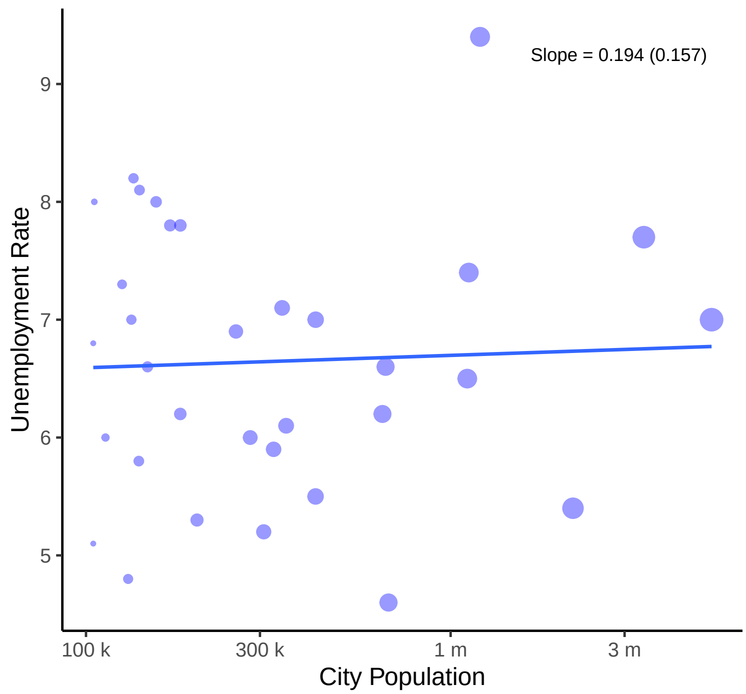
Notes: See Section 2 for a description of the sample construction procedure. Big city refers to the largest three cities in Canada, and small city refers to the remaining ones.

Table A.2: Urban earnings growth premium by sector

Sector	Total	Between	Within
Construction	0.166^{*}	0.078^{*}	0.088
	(0.094)	(0.046)	(0.056)
Manufacturing	0.133^{**}	0.080^{***}	0.053
	(0.058)	(0.027)	(0.047)
Retail and wholesale	0.074^{**}	0.122^{***}	-0.048^{**}
	(0.035)	(0.020)	(0.022)
Transportation	0.259^{***}	0.092^{**}	0.166^{***}
	(0.074)	(0.038)	(0.056)
Tradeable services	0.163^{***}	0.095^{***}	0.068^{**}
	(0.035)	(0.008)	(0.034)
Admin, education and health	0.179^{**}	0.050	0.129^{**}
	(0.068)	(0.053)	(0.052)
Entertainment and hospitality	0.121^{**}	0.103^{***}	0.018
	(0.060)	(0.026)	(0.042)

Notes: The regressions are weighted by the number of worker-year observations in each city-sector bin. Robust standard error clustered at the city-sector level. Significance level: *p<0.1; **p<0.05; ***p<0.01

Figure A.1: Unemployment rate and city population



Notes: The figure plots the unemployment rate against the city population for 33 Census Metropolitan Areas of 2016. Data source: StatsCan Table 14-10-0096-01. The point estimate with its standard error of the population-weighted OLS regression is reported in the figure.