

Canadian Labour Economics Forum

WORKING PAPER SERIES

Labour Market Power, Firm Productivity, and the Immigrant-Native Pay Gap

Stephen Tino
(Toronto Metropolitan University)

CLEF WP #88 (updated 2026-09-19)

Labor Market Power, Firm Productivity, and the Immigrant-Native Pay Gap*

Stephen Tino†

September 19, 2026

Click here for latest version

Abstract

How much of the immigrant–native pay gap reflects employer wage-setting power rather than productivity? Using matched employer–employee data covering all Canadian tax filers, I estimate an oligopsony model in which labor supply elasticities vary by firm and immigrant status, and production technology varies within and across firms. Employers pay immigrants 23% below their marginal revenue product and natives 16% below, while marginal products differ by 5%. Counterfactuals show that monopsony power explains 81% of the gap, and that it also concentrates immigrants in lower productivity matches, so reducing it would raise aggregate output as well as immigrant pay.

JEL Classification Numbers: J01, J15, J23, J31, J42

*I am extremely indebted to Kory Kroft, Ismael Mourifié, and Carolina Arteaga for unwavering guidance and support. I also thank Abdelrahman Amer, Ben Couillard, Benoit Dostie, David Green, Sam Gyetvay, Jeff Hicks, Billy Huang, Alexandre Lehoux, Matthew Notowidigdo, Mikal Skuterud, Román Zárate, Christopher Worswick, participants at the CEA Annual Conference 2024, participants at the CLEF Annual Conference 2025, participants in the University of Toronto SWEAT and Empirical Microeconomics seminar series, participants in the S4 Structural Reading Group, and the Toronto RDC staff. First version: Oct 29, 2024.

†Department of Economics, Toronto Metropolitan University (email:stino@torontomu.ca)

1 Introduction

In all 33 high-income countries surveyed by the International Labour Organization of the United Nations, immigrants tend to earn less than native-born workers, with an average pay gap of 13% across countries (Amo-Agyei 2020). Canada, which has the highest proportion of immigrants among G7 countries, is no exception, despite having an immigration policy that is explicitly designed to attract high-skilled workers.[1] According to the 2016 Canadian Census, immigrants earn 21% less than natives on average, and this pay gap *widens* to 35% when controlling for education and experience.

The literature offers several explanations for the immigrant-native pay gap, including differences in language skills (Chiswick and Miller 1995), literacy (Ferrer, Green, and Riddell 2006), quality of schooling (Fortin, Lemieux, and Torres 2016), discrimination (Oreopoulos 2011), and job mobility (Pendakur and Woodcock 2010; Skuterud and Su 2012; Javdani and McGee 2018). Recent studies emphasize the role of employer wage-setting policies, using two-way fixed effects models in the Abowd, Kramarz, and Margolis (1999, hereafter AKM) tradition to estimate firm pay premia and assess their contribution to the pay gap.

These studies show that immigrants are disproportionately employed in "low-wage firms," i.e., firms with relatively low pay premia (Damas de Matos 2017; Dostie et al. 2023; Arellano-Bover and San 2024; Amior and Stuhler 2024; Gyetvay and Keita 2024; Lehrer and Rawling 2025). However, firm pay premia arise from several distinct factors, including firm productivity, compensating differentials, and labor market power (Card et al. 2018), and we do not know the relative contribution of each. The existing literature provides little insight into the relative contribution of these factors due to the methodological challenges associated with measuring them and the substantial data requirements involved.

In this paper, I investigate the importance of labor market power and firm productivity for understanding the immigrant-native pay gap. The main data source is the Canadian Employer-Employee Dynamics Database (CEEDD), a matched employer-employee dataset covering 100% of tax filers from 2001 to 2019, which I also supplement with data from the 2016 Canadian Census. My empirical setting is the near universe of private-sector employment in Canada, comparing prime-age, full-time workers born in Canada to workers who hold or have ever held permanent residency. The analysis proceeds in two parts. I first document a set of descriptive facts about who immigrants work for, what those employers produce, and what they pay, which together point toward monopsony power without being able to quantify it precisely. I then

1. Due to recent record-breaking growth in immigration, roughly one-quarter of individuals in Canada are immigrants (Statistics Canada 2022). Moreover, a key feature of Canada's immigration policy is the point system that selects applicants with high levels of human capital (Beach, Green, and Worswick 2011). That policy treats the gap as a skill problem, and so implicitly locates its cause in the characteristics immigrants arrive with rather than in the market that hires them.

estimate a structural model of an imperfectly competitive labor market, which separates what a worker produces at a given employer from what that employer pays her relative to it, and use the model to decompose the pay gap in a general equilibrium framework. Each channel in that decomposition corresponds to a policy lever: the skill mix of immigrants and the technology they use in production, the frictions that tie them to particular employers, and the concentration of labor markets. General equilibrium matters because reducing monopsony power also changes where immigrants work.

I begin by asking whether the gap is reducible to differences in observable skill. For this, I turn to the 2016 Census, as the matched employer-employee data is derived from tax records and lacks comprehensive information on education or occupation. Despite an unconditional pay gap of 20.8 log points, immigrants are considerably more educated, with 43.8 percent holding at least a bachelor's degree versus 20.1 percent of natives. Controlling for education and experience in a Mincer-style regression *increases* the gap to 35.5 log points, and it remains at 27.3 log points with fully interacted industry-by-occupation fixed effects. The widening is driven entirely by the economic class—immigrants admitted through a points system that scores education, experience, and language, and therefore the group whose credentials should be most comparable to natives'—whose unconditional gap of 9.6 log points triples once measured skill is held constant. Either observed skill in the Census is a poor proxy for actual productive skill, or the conditional gap is a wedge between pay and productivity rather than a difference in productivity.

This shifts the focus to employer heterogeneity, using the matched employer-employee data. Immigrants are strikingly segregated across firms: the average immigrant works at a firm that is 51.0 percent immigrant, against 11.6 percent for the average native. Estimating an AKM model in these data, I confirm what Dostie et al. (2023) and the broader two-way fixed effects literature find: immigrant-intensive firms pay premia that are 8.9 log points lower, with the entire distribution shifted left. What is different about these firms? Part of the answer is productivity. Weighting by employment, firms that hire immigrants generate 4.8 log points less revenue per worker than those that hire natives, consistent with Åslund et al. (2025). But the estimated pass-through of revenue per worker into firm pay premia is only about 0.1, so this difference can account for only a small fraction of the pay gap. Productivity is therefore unlikely to be the whole story, and something else must differentiate these employers.[2]

The natural candidate is monopsony power, and two facts point toward it. First, immigrants are 33 percent more exposed to employer concentration, driven by geographic clustering into a small number of cities rather than by working in thin local markets. Second, and more directly, comparing firms that are *equally productive* but differ in how intensively they hire immigrants,

2. These estimates concern the pass-through of *average* product to firm pay premia, whereas wages are set against *marginal* products. The exercise is suggestive rather than conclusive; marginal products are estimated structurally in Section 5.

pay premia at immigrant-intensive firms are roughly 6.8 log points lower at every point of the revenue-per-worker distribution. Employers appear to retain a larger share of what is produced when the workers producing it are immigrants.

Each of these comparisons is descriptive, and the AKM firm premium at the centre of them is a statistical object rather than a structural one. Because firm effects are identified from movers, any component of pay common to all of a worker's employers is absorbed into the worker effect, where a level of wage-setting power constant across the firms hiring immigrants is indistinguishable from human capital. Evidence of this kind can establish that immigrant-intensive firms pay less than their productivity predicts, but it cannot quantify how much of the gap is wage-setting power and how much is productivity. That calls for a structural model.

I therefore estimate a wage-posting oligopsony model adapted from Chan et al. (2024, hereafter CKMM). Its central object is the labor supply curve facing an individual employer, which I estimate *separately for immigrants and natives*.[3] Two forces give employers wage-setting power in the model: labor market concentration and the substitutability of employers. Firms also vary in total factor productivity, returns to scale, and match quality, allowing the data to distinguish wage-setting power from productivity as explanations for the pay gap. Immigrants and natives are imperfect substitutes in production, so a firm's demand for one group depends on its employment of the other. The equilibrium wage of a worker type is the marginal revenue product of labor (MRPL) multiplied by a markdown determined entirely by the labor supply elasticity.

The estimates imply that employers withhold 17.6 percent of the MRPL on average, with heterogeneity across immigrant status: 16.4 percent from natives compared to 22.8 percent from immigrants. Across admission classes the markdowns are 20.6 percent for economic-class immigrants, 24.6 percent for the family class, and 27.0 percent for refugees, reproducing the ranking of the conditional pay gaps in the Census and tracking the degree to which admission is conditioned on labor market attachment. Moreover, the model's welfare-relevant measure of exposure to labor market concentration, the generalized concentration index (GCI), indicates that immigrants' exposure is almost twice as large as natives'. Against all of this, the gap in MRPL is 5.3 percent, consistent with the descriptive facts that suggest differences in productivity are modest.

To decompose the pay gap, I conduct counterfactual analyses that impose restrictions on the model—either equalizing primitives across groups or altering the wage-setting rule—and re-solve for the equilibrium, letting wages, markdowns, and the allocation of workers across firms adjust. I group the restrictions into four blocks, each corresponding to a distinct policy lever: the substitutability of employers, which is acted on by policies that improve information and mobility; market structure, the object of antitrust enforcement; the value of non-employment, which de-

3. CKMM study the sources of overall wage inequality rather than a pay gap between two groups, and do not consider immigration.

pends on settlement supports and social assistance; and the skill mix of the workforce together with production technology, which is determined by immigrant selection criteria, training, and industrial policy. The difficulty is that the contributions of these blocks are not additive: because the channels interact, what any one block contributes depends on which others are already imposed. I therefore report the Shapley value of each block (Shorrocks 2013), which averages its marginal contribution over every order in which the restrictions can be imposed.

Monopsony power accounts for 80.8 percent of the gap, of which 74.6 points are attributable to the substitutability of employers and 6.3 to differential exposure to concentration. The value of non-employment accounts for 7.9 percent and skill mix and technology for 11.3 percent. Furthermore, the counterfactuals show that employer wage-setting power leads to misallocation that concentrates immigrants in lower productivity matches. Ultimately, the results speak to the importance of designing policy that can improve immigrants' outside options and reduce employer wage-setting power. Because wage-setting power also concentrates immigrants in less productive matches, reducing these frictions would not be purely redistributive: it would move immigrants toward better matches, and aggregate output would rise along with immigrant pay.

This paper contributes to the literature on monopsony power and immigration (Amior and Manning 2020; Naidu, Nyarko, and Wang 2016; Depew, Norlander, and Sorensen 2017; Dostie et al. 2023; Amior and Stuhler 2024; Wang 2021; Hirsch and Jahn 2015; Kroft et al. 2025). The closest antecedent is Hirsch and Jahn (2015), who apply the dynamic monopsony framework of Manning (2003) to German data and find that firms mark immigrant wages down more than natives'. I estimate a markdown differential of 6.4 percentage points, close to their 7.7 log points. This agreement is reassuring, and my paper differs in three important respects. First, marginal revenue products are estimated, so the markdown is measured against productivity, and differences in productivity are quantified as a rival explanation for the gap. Second, wage-setting power has two sources, labor market concentration and the substitutability of employers, so my paper can shed some light on why firms mark down wages more for immigrants. Third, the importance of wage-setting power is quantified in a general equilibrium framework and thus can speak directly to what would happen to the pay gap if government policies that reduced wage-setting power were implemented. Furthermore, I find that employer wage-setting power clusters permanent residents in low-wage, low-productivity matches, complementing the results in Kroft et al. (2025) showing that relaxing mobility constraints allows temporary foreign workers to move to more productive and higher paying firms.

The paper also contributes to the literature that decomposes pay gaps between groups. The seminal approach separates the gap into a part explained by observed worker characteristics and a residual (Oaxaca 1973; Blinder 1973). Recent work adds employers to that accounting, using two-way fixed effects models to split the gap into a component operating through worker skill

and a component operating through the firms that each group works for (Card, Cardoso, and Kline 2016; Dostie et al. 2023; Arellano-Bover and San 2024; Amior and Stuhler 2024; Gerard et al. 2021). These decompositions establish that where a group works matters, but the firm component is a statistical residual, defined holding the allocation of workers across firms fixed. In contrast, I apply the CKMM framework to develop a new decomposition of a pay gap into employer wage-setting power and productivity, computed by re-solving the model for equilibrium and letting workers reallocate, so that the results can be mapped to policy levers.

The remainder of the paper proceeds as follows. Section 2 describes the CEEDD, the population I study, and the construction of the sample. Section 3 presents the descriptive evidence on the pay gap, employer segregation, firm productivity, and labor market concentration, and draws out the features a model would need to account for them. Section 4 develops the model and discusses identification and estimation. Section 5 reports the estimated elasticities, markdowns, and technology parameters. Section 6 decomposes the pay gap, and Section 7 concludes.

2 Data

2.1 The AKM Sample

I begin with the full matched employer-employee data covering 2001 to 2019 and follow closely the sample restrictions outlined in Dostie et al. (2023).[4] The resulting sample is the near-universe of Canadian private-sector employment relationships for full-time, prime-age workers that are either born in Canada or immigrants that have obtained permanent residency.[5]

Information on immigrant status is obtained from the Longitudinal Immigration Database (IMDB), and I classify individuals as permanent residents in all years after their landing year, excluding foreign-born individuals whose landing year is missing.[6] Permanent residents are restricted to those admitted through one of the three main classes: the *Economic Class*, the *Family Class*, or the *Refugee and Humanitarian Class*.[7] Temporary residents are excluded, since temporary permits carry legal restrictions on employer mobility that require separate treatment.

4. I depart from their procedures only in the definition of full-time equivalent (FTE) workers, which I obtain from Li, Dostie, and Simard-Duplain (2023). I choose this definition because it is calculated based on the annual earnings that a worker would receive if they work full-time at the minimum wage for the entire year.

5. In Section 2.3, I compare the summary statistics for this "AKM sample" to summary statistics for full-time workers in the private sector in the 2016 Canadian Census, and find that the two samples are quite similar.

6. Permanent residents may live and work anywhere in the country indefinitely and differ from citizens only in a narrow set of rights, including eligibility to run for political office and hold certain government jobs requiring high-level security clearance. I thus do not distinguish permanent residents who later obtain citizenship.

7. These classes differ in how tightly admission is conditioned on labor market attachment, which turns out to be the dimension along which the pay gap varies most. Appendix C describes the admission categories in more detail.

Individuals with missing marital status, those who do not identify as male or female, and those outside the working age of 25 to 59 are excluded. Furthermore, the sample is limited to individuals whose employment income is at least as large as their self-employment income, where self-employment income includes earnings from business, farming, fishing, rental, commissions, and professional activities.

Firms in the public sector (NAICS 91), education (NAICS 61), and health sectors (NAICS 62) are excluded. The sample is also restricted to incorporated firms that meet several criteria: they must have at least $50,000 in revenue, at least $100 in value added per worker, and revenue that is at least as large as the total wage bill, where all monetary thresholds are expressed in 2012 dollars. Additionally, these firms must have at least two employees, where employment is defined as the average of all non-zero monthly employment submissions from the PD7.

Since the CEEDD derives its data from tax records, it lacks information on hours worked. To address this, the sample is restricted to full-time equivalent (FTE) workers, defined as those earning at least approximately $18,000 in 2012 dollars, which is calculated by multiplying the minimum wage by an average full-time work schedule of 38.8 hours per week over 48 weeks (Li, Dostie, and Simard-Duplain 2023). Moreover, individuals in the CEEDD may have multiple T4 records if they hold multiple jobs. I restrict the data to each individual's primary job, defined as the job that provides the highest earnings in any given year.

2.2 Labor Markets and Local Employers

Labor markets are defined following Lamadon, Mogstad, and Setzler (2022, hereafter LMS) as combinations of two-digit NAICS codes and geographic locations. Geographic locations are based on Census Metropolitan Areas (CMAs) or Census Agglomerations (CAs) as defined in the 2016 Census. CMAs and CAs consist of population centers and adjacent municipalities with high commuting flows, resembling commuting zones in the United States. Labor markets in the territories (Yukon, Northwest Territories, and Nunavut) are excluded from the analysis.

In the CEEDD, both worker and firm locations are available. Worker location is derived from the T1PMF, while firm location comes from the NALMF. However, firms in the CEEDD are defined by their Enterprise ID in the Business Registry for tax purposes, which means the "firm location" corresponds to the headquarters. For multi-location firms, I treat each Enterprise ID-location pair as an independent unit, where the "location" of the unit corresponds to that of the workers. I impute revenue for each of these local units by allocating the observed revenue associated with the Enterprise ID based on the local unit's share of the total wage bill, following CKMM. In my analyses, I estimate distinct production technology parameters for each local unit and refer to these local units as "firms." Importantly, however, I use the aggregate revenue associated with the Enterprise ID for the identification of the labor supply parameters (Section 4.5).

2.3 Descriptive statistics

Table 1 reports summary statistics for the AKM sample; the corresponding moments are close to those in Dostie et al. (2023). The sample contains 10.3 million native-born workers and 2.95 million individuals who hold or have held permanent residency, observed over 74.5 and 17.6 million person-years respectively, so immigrants make up roughly 22% of the workers in the sample and 19% of the person-years. The two groups are demographically similar on the dimensions the tax data records: the male share is 65.0% for natives and 61.5% for immigrants, and mean age is 42.1 against 41.7. They differ sharply in geography, with 54.5% of immigrants in Ontario against 35.9% of natives, and 13.6% in Quebec against 27.1%.

Mean annual earnings are roughly $54,500 for immigrants against $65,500 for natives, a difference of 16.8%, which corresponds to a pay gap of 18.4 log points. Immigrants also work at larger firms. The median immigrant's employer has 75 employees against 29 for natives, and mean log firm revenue is 17.24 against 16.51.

Comparison with the Census

A natural question is whether the AKM sample remains representative of prime-age full-time employment in Canada, given that full-time status is imputed from an earnings threshold rather than observed. Appendix Table A1 answers it using the 2016 Census, where full-time employment status is reported directly and the same age and sector restrictions can be applied. Because the IMDB records landings only from 1980 onward, the immigrants in the AKM sample are necessarily post-1980 arrivals, and I therefore restrict the Census sample to immigrants admitted in 1980 or later so that the two samples cover the same cohorts.

The two samples line up closely. The male share is 64.4% for natives and 60.4% for immigrants in the Census, against 65.0% and 61.5% in the AKM sample. The Census age bands imply a mean age close to 42 in both groups, against 42.1 and 41.7. The provincial distribution agrees to within roughly three percentage points for Ontario, Quebec, and British Columbia in both groups, with the largest discrepancy the immigrant share in Ontario, 51.1% against 54.5%.

Earnings are also similar across the two samples. The immigrant-native pay gap is 20.8 log points in the Census versus 18.4 in the AKM sample. Appendix Table A2 shows the same agreement by admission category: Census gaps relative to natives of 9.6 for economic immigrants, 35.0 for the family class, and 33.3 for refugees, against 8.3, 31.5, and 32.6 in the AKM sample.

3 Stylized Facts

This section establishes five facts about the immigrant–native pay gap in Canada. These facts are consistent with employers having more monopsony power over immigrants compared to native workers and motivate the structural model described in Section 4.

Fact 1 shows that the gap is not accounted for by measured skill, industry, or occupation: conditioning on these characteristics widens the gap rather than closing it. Fact 2 moves the object of interest from the worker to the employer. Immigrants are segregated into a distinct set of firms, and those firms pay systematically lower wage premia. This naturally raises the question that organizes the remainder of the section: what is different about firms that hire immigrants?

Facts 3–5 point to an answer. Immigrant-intensive firms generate less revenue per worker, but the difference is modest (Fact 3). Immigrants' exposure to labor market concentration is high once geographic clustering is taken into account (Fact 4). And conditional on productivity, immigrant-intensive firms pay lower wage premia at every point of the productivity distribution, with the productivity difference accounting for only about a quarter of the premium gap (Fact 5).

Fact 1: Conditioning on observable skill widens the pay gap

I begin outside the CEEDD with the 2016 Canadian Census, which, unlike the tax data, reports education, official-language ability, and occupation for both immigrants and natives. Table 2 reports Mincer wage regressions for full-time workers aged 25 to 59, excluding the education, health, and public administration sectors. Column 1 establishes a baseline: the unconditional immigrant-native pay gap is 20.8 log points. If this gap were driven by a deficit in observable skills, controlling for those skills should reduce it. Instead, the gap widens by more than two-thirds. Column 2 adds age, gender, education, primary official language, and marital status, pushing the immigrant penalty to 35.5 log points. Columns 3 through 5 incorporate industry and occupation fixed effects. While these absorb some of the penalty, the gap remains large in magnitude: controlling for interacted industry-by-occupation, immigrants still earn 27.3 log points less than natives. Paradoxically, immigrants and natives with identical credentials working in the exact same jobs are further apart in pay than those chosen at random.

The pattern is not one of immigrants holding weaker credentials. In the Census sample, 43.8% of immigrants hold at least a bachelor's degree, against 20.1% of natives. Table A3 shows that the widening is driven entirely by the "economic class"—immigrants admitted via a points system that heavily weights human capital. Because this group is intensely screened, with 60.0% holding at least a bachelor's degree, their unconditional pay gap is just 9.6 log points, compared to 35.0 log points for the family class and 33.3 for refugees. Yet holding measured skill constant more than triples the economic-class penalty, to 33.9 log points (Column 2), while the gaps for the

other classes remain similar. Once industry and occupation are fully interacted, the three classes converge to between 25.4 and 28.4 log points.

There are two ways to read this pattern. The first is that the Census records formal credentials rather than productive skill, and that human capital acquired abroad is imperfectly portable to the Canadian labor market. A degree obtained in another country may simply be worth less, for example because the tasks of the job are different or because occupational licensing bodies do not recognize foreign certification.

The second reading is that the widening has little to do with human capital at all. Suppose the skill transfers intact, and the immigrant and the native are equally productive in the same job. They may still be paid differently if their employers engage in monopsonistic wage discrimination. Under this reading the conditional gap is a wedge between pay and productivity rather than a difference in productivity.

Mincer regressions in the Census cannot separate these two accounts; the data record credentials and not productivity, and do not include information on employers. In what follows, I turn to the matched employer–employee data, and begin by asking how much of the gap is attributable to *where* immigrants work.

Fact 2: Immigrant-intensive firms pay lower wage premia

How do employer wage-setting policies contribute to the immigrant-native pay gap? I begin by documenting the segregation of immigrants across employers in the Canadian labor market. Table 1 reports that the average immigrant works at a firm where 51.0% of employees are immigrants, while the average native works at a firm where the share is 11.6%, suggesting a significant degree of segregation.

Is this segregation relevant for the immigrant-native pay gap? Several recent papers estimating two-way fixed effects models of log earnings in the tradition of AKM (Abowd, Kramarz, and Margolis 1999) find that immigrants sort into employers that pay low wage premia. Dostie et al. (2023), using the same Canadian data as this paper, reach this conclusion; so do Damas de Matos (2017) for Portugal, Arellano-Bover and San (2024) for Spain, and Amior and Stuhler (2024) and Gyetvay and Keita (2024) for Germany.

I confirm this pattern in the AKM sample by estimating a two-way fixed effects model of log earnings and investigating how firm pay premia differ across employers of immigrants and natives.[8] Figure 1 focuses on firms that hire at least one immigrant, splits these firms into two

8. Estimating this two-way fixed effects model requires one additional restriction beyond those outlined in Section 2.1: to identify the firm effects, the sample is restricted to the largest connected set of employers. See Appendix D for more details on the estimation of the AKM model.

groups—*native-intensive* and *immigrant-intensive*—based on whether their time-averaged immigrant share falls below or above the median, and plots the distribution of estimated firm effects for each group. The mean firm effect is −0.199 at native-intensive firms and −0.288 at immigrant-intensive firms, a difference of 8.9 log points. The entire distribution for immigrant-intensive firms is shifted to the left, indicating that this gap is not merely driven by extreme values.[9]

Notably, an estimated firm pay premium is a statistical object, not a structural one. The two-way fixed effects decomposition is structurally silent about the *level* of wage-setting power. Suppose every employer marks immigrant wages down by a common factor. That markdown is constant across all of an individual's employers in the data, so it is absorbed into the worker effect α_i, where it is observationally identical to lower human capital. Only variation in markdowns *across* employers can be captured by firm effects. Appendix D.3 makes this point more formally.

Nonetheless, the AKM decomposition shows that firm heterogeneity contributes to the immigrant-native pay gap. Next, I explore whether this arises because immigrants sort into firms that produce less per worker.

Fact 3: Immigrant-intensive firms produce less per worker, but the difference is modest

Under any wage-setting model in which workers receive a share of the surplus their employer generates—rent sharing, bargaining, or wage posting against an upward-sloping supply curve—a firm that produces less per worker pays less. It is therefore possible that the premium gap in Fact 2 is ultimately reducible to a gap in employer productivity.

To begin, I rank all firms in the AKM sample by revenue per worker, take the subset of firms that hire at least one immigrant, split them into immigrant-intensive and native-intensive firms, and plot the distribution of those ranks across the two groups.[10] The results are presented in Figure 2. The density for immigrant-intensive firms is highest in the bottom quintile and declines through the middle of the distribution before flattening, while the density for native-intensive firms rises steadily and peaks at the top. The mean percentile is 45.9 among immigrant-intensive firms and 58.0 among native-intensive firms, and the separation is a shift of the entire distribution rather than a feature of either tail.

9. Note that I estimate a *single* firm effect ψ_j for each firm, common to immigrants and natives, and compare firms according to their immigrant employment share. I do not estimate separate firm effects by immigrant status. Comparing ψ_j^{imm} to ψ_j^{nat} requires an additional assumption, because each set of firm effects is identified only up to a group-specific normalization and a difference in levels between them is not interpretable without pinning the two normalizations to a common scale. Dostie et al. (2023) do estimate immigrant-specific firm effects in these data, invoke such an assumption, and find that immigrants and natives at the same employer receive very similar pay premia, justifying my approach.

10. The sample and the split into immigrant- versus native-intensive are exactly those of Figure 1.

A more natural analog to the immigrant–native pay gap is the difference computed over the full AKM sample and weighted by the number of workers at each firm, which is also reported in Table 1. Measured this way, immigrants work at firms whose mean log revenue per worker is 4.8 log points lower than natives', with the medians differing by 5.6.

Immigrants therefore do work at firms that generate less revenue per worker, but the individual-weighted difference is modest relative to the pay gap, and only a fraction of it would be passed through to wages.[11] This shifts attention from what these firms produce to what they are able to retain, and the natural starting place is to investigate differences in wage-setting power. Fact 4 takes up its conventional proxy: labor market concentration.

Fact 4: Immigrants are more exposed to labor market concentration

The conventional first-order proxy for labor market power is concentration. Index worker groups by k (here, immigrant or native-born), firms by j, and labor markets by $g \in \mathcal{G}$, where a market is a two-digit NAICS by CMA/CA cell as defined in Section 2.2. Let n_{kj} be the number of type-k workers at firm j, let $N_{kg} = \sum_{j \in \mathcal{J}_g} n_{kj}$ be the number in market g, and let $N_k = \sum_g N_{kg}$ be the total. The standard measure is the Herfindahl-Hirschman Index (HHI) of type-k employment across the firms in a market,

$$H_{kg} \equiv \sum_{j \in \mathcal{J}_g} \left(\frac{n_{kj}}{N_{kg}} \right)^2, \tag{1}$$

summarized for a group by the employment-weighted average $H_k^W \equiv \sum_g s_{kg} H_{kg}$, where $s_{kg} = N_{kg}/N_k$.

An average of within-market indices, however, does not capture concentration *between* markets. What disciplines a firm's wage-setting power is the set of employers a worker would actually move to, and that set is restricted by being confined to few markets exactly as it is by facing few firms inside one. This matters here because immigrants in Canada cluster geographically: although immigrants are roughly one quarter of Canada's population, more than half live in Toronto, Montreal, or Vancouver (Statistics Canada 2022), and Table 1 shows 54.5% of immigrants in Ontario against 35.9% of natives. This is a general feature of immigrant settlement rather than a Canadian peculiarity. Albert and Monras (2022) document that immigrants in the United States concentrate disproportionately in a small number of large, expensive metropolitan areas, Card (2001) shows that this clustering is stable enough to serve as an instrument for local immigrant inflows, and Edin, Fredriksson, and Åslund (2003) and Damm (2009) document the same enclave formation in Sweden and Denmark.

Two features of the HHI make it a poor basis for comparing immigrants' and natives' exposure

11. Fact 5 quantifies how much of the premium gap the observed difference in revenue per worker can account for.

across both margins. First, it is mechanically inflated for smaller groups, because H_{kg} is the probability that two type-k workers drawn at random *with* replacement work at the same firm and so counts the event in which the same worker is drawn twice. Immigrants are roughly a fifth of employment in my sample and are absent from many markets altogether, so their market-level cell counts are systematically smaller than natives'. Second, the overall index $H_k \equiv \sum_{j\in\mathcal{J}}(n_{kj}/N_k)^2$ does not factor as $H_k = H_k^B \times H_k^W$, where H_k^B is the between-market analog of Equation (1), so the HHI admits no clean decomposition into the two margins.

Both problems are solved by working with the collision probability: the probability that two type-k workers, drawn at random *without* replacement, are employed by the same firm,

$$C_k \;\equiv\; \sum_{j\in\mathcal{J}} \frac{n_{kj}\,(n_{kj}-1)}{N_k\,(N_k-1)}. \tag{2}$$

Because it is defined without replacement, it is comparable across groups of different size. And because a pair collides at a firm if and only if it is in the same market *and* collides within it, splitting Equation (2) by market gives the exact multiplicative decomposition

$$C_k \;=\; \underbrace{C_k^B}_{\text{between}} \times \underbrace{\sum_{g\in\mathcal{G}} \pi_{kg}\, C_{kg}}_{\equiv\, C_k^W \text{ (within)}}, \qquad \pi_{kg} \equiv \frac{N_{kg}(N_{kg}-1)}{N_k(N_k-1)\,C_k^B}, \tag{3}$$

where C_k^B is the probability that two randomly drawn type-k workers are employed in the same market, C_{kg} is the within-market collision probability, and $\sum_g \pi_{kg} = 1$ by construction. Appendix E derives Equation (3) and gives the exact size correction relating C_{kg} to H_{kg}. Both factors in Equation (3) are probabilities, so the two margins are on a common scale.

Panel A of Table 3 reports the three collision probabilities. *Within* labor markets, immigrants face *less* concentration than natives: their within-market collision probability is 0.45 times that of natives. *Between* labor markets, they face nearly three times the concentration natives do, a ratio of 2.97. The between-market component dominates, so overall, immigrants' exposure to labor market concentration is 1.33 times that of natives. Panel B reports the HHI analogs. The size correction matters for the within-market comparison, where the HHI ratio is 0.75 rather than 0.45, but it leaves the headline unchanged: the overall HHI is 1.34 times larger for immigrants.

I conclude that immigrants are more exposed to labor market concentration than natives, which points to monopsony power as a possible explanation for the premium gap in Fact 2. However, concentration is only one indicator of monopsony. More broadly, employers can mark down wages whenever their workers' outside options are limited, i.e., whenever workers offered a wage below their marginal product cannot easily take a job elsewhere. Such monopsony power is pervasive (Manning 2021; Card 2022), even though it may not be reflected in employment shares.

This suggests the next exercise. If the workers at immigrant-intensive firms have more limited outside options—for whatever combination of reasons—those firms should be able to retain a larger share of what is produced, and so should pay less for a given level of productivity.

Fact 5: Conditional on productivity, immigrant-intensive firms still pay less

Fact 3 showed that immigrant-intensive firms produce modestly less per worker, and Fact 4 showed that immigrant labor markets are more concentrated. Next, I test whether monopsony power contributes to the immigrant-native pay gap more directly by investigating whether firms are able to capture more of what is produced from immigrants relative to natives, without taking a stand on the source of the monopsony power that allows them to do so.

The test follows the logic of Card et al. (2018) and Beauregard et al. (2025). Intuitively, comparing firm effects *conditional on productivity* should be informative about the wedge between what is produced and what is paid to workers.[12] If immigrant-intensive firms pay less only because they are less productive, the gap should vanish once productivity is held fixed. If a gap survives at every level of productivity, immigrant-intensive firms are passing on a smaller share of what they produce, suggesting a larger wedge between productivity and wages.

The exercise uses the same sample as Figures 1 and 2: firms employing at least one immigrant in the estimation panel, which is the set over which the immigrant-intensity split is defined. Firms are grouped into bins of log revenue per worker and the mean firm effect is plotted against the bin mean, separately for immigrant- and native-intensive firms.[13] Figure 3 presents the results, and two features stand out.

First, the relationship is positive for both groups, with a slope of roughly 0.08. More productive firms pay higher premia, as any pass-through model predicts. Both relationships flatten slightly toward the upper end of the productivity distribution, a little more so for immigrant-intensive firms; the estimated slopes are 0.080 for native- and 0.074 for immigrant-intensive firms.

Second, the two lines are separated by a persistent vertical gap. Conditional on log revenue per worker, immigrant-intensive firms pay approximately 6.8 log points less than native-intensive firms. Since the unconditional premium gap is 8.9 log points, differences in revenue per worker account for roughly a quarter of it, and three quarters survives.

12. To be precise, revenue per worker is about *average* product, and wages are set based on *marginal* product, and with diminishing returns to scale, the two do not agree. Nonetheless, this test is useful as a descriptive statistic to motivate the model in Section 4, which estimates the marginal revenue product and the wedge between the marginal revenue product and wages rigorously.

13. Appendix Figure A1 repeats the exercise described in this section using value added per worker, and the results are similar. Revenue per worker is the preferred measure here because it maps more directly to the structural model in Section 4.

Taken together, these facts point to the importance of labor market power and firm productivity for understanding why immigrants earn less, and they suggest that monopsony power is an important factor. However, these stylized facts are mainly suggestive and motivate further investigation. In the next section, I describe the structural model that can accurately disentangle between various causal mechanisms and quantify precisely the contribution of labor market power and firm productivity to the immigrant-native pay gap.

4 Model

4.1 What kind of model do the facts call for?

Section 3 suggests the model should include the following three characteristics:

(i) Wage-setting power, with firm-specific supply elasticities estimated by group. Fact 5 is consistent with employers paying immigrants a smaller share of what they produce. Testing that requires a model in which firms face upward-sloping labor supply curves, and in which the elasticity of that curve is estimated separately for each worker type.

(ii) Multiple sources of that power. Immigrants' overall exposure to labor market concentration is higher, but this gap is entirely between markets; within their markets they face *less* concentration than natives (Fact 4). The model must therefore allow wage-setting power to arise from concentration both within and between labor markets. It must also accommodate sources of wage-setting power that operate independently of market structure, as *any* mechanism that limits the outside options available to a worker at a given level of concentration can theoretically lead to wage-setting power.

(iii) Flexible productivity heterogeneity, across firms and within them. Fact 3 showed that immigrant-intensive firms generate less revenue per worker, and Fact 5 that pay premia rise with productivity, so the model must include firm-level productivity and allow it to pass through to wages. Fact 1 adds a second margin: conditioning on credentials widened the gap, and one reading is that foreign human capital is imperfectly portable, which is heterogeneity in productivity *within* the firm, across worker types.

These requirements are met jointly by the oligopsonistic wage-posting model described next.

4.2 Setup

Workers are grouped into discrete types, and I extend CKMM by including immigrants in that classification. Permanent residency in Canada is granted under three categories—*Economic*, *Family*, and *Humanitarian and Refugee* (Appendix C)—and Fact 1 showed that the gap behaves very differently across them, so this is an important dimension of worker heterogeneity to include.

Workers are also classified by gender, following the labor market power and gender pay gap literature (Sharma 2024; Robinson 1933; Webber 2016), and by continent of origin, to allow for heterogeneous returns to education (Fortin, Lemieux, and Torres 2016; Dostie et al. 2023).[14]

Each worker i has type $k \in \mathcal{K}$; there are M_t workers in period t and m_{kt} of each type, with $\sum_{k=1}^{K} m_{kt} = M_t$. There are $g \in \mathcal{G}$ local labor markets, each defined by location and industry (Section 2), and J firms indexed by $j \in \mathcal{J}$, with $\mathcal{J}_g$ the firms in market g.

4.3 Labor supply

The object of interest is the supply curve facing an *individual employer*: the type-k labor a firm can hire as a function of the wage it posts, holding other firms' wages fixed. I microfound it with an additive random utility model (ARUM), following CKMM. The class is general—it nests LMS, Berger, Herkenhoff, and Mongey (2022), and CKMM—and it is agnostic about the frictions in requirement (ii), which the non-wage terms absorb.

Assumption 4.1 (Worker choice)**.** *Workers of type k in market g choose among the employers of that market, $j \in \mathcal{J}_g$, and non-employment, indexed 0, to maximize a payoff additive in (i) the log wage $\log w_{kjt}$, with type-specific weight $\beta_k > 0$, where w_{k0t} is the value of non-employment; (ii) a component ξ_{kjt} common to all workers of type k facing firm j in period t; and (iii) an idiosyncratic component drawn from a generalized extreme value distribution whose nests are local labor markets, with within-nest dispersion governed by $\sigma_{kg} \geq 1$.*

Let s_{kjt} denote the share of type-k workers employed at firm j, s_{k0t} the share non-employed, and $s_{kjt|g} \equiv s_{kjt} / \sum_{h \in \mathcal{J}_g} s_{kht}$ firm j's share of type-k employment within its market.

Lemma 4.1 (Firm-level labor supply)**.** *Under Assumption 4.1, equilibrium employment shares satisfy*

$$\underbrace{\frac{s_{kjt}}{s_{k0t}}}_{\text{quantity supplied to } j} = \underbrace{\exp(\xi_{kjt})}_{\text{intercept}} \times \underbrace{\left(\frac{w_{kjt}}{w_{k0t}}\right)^{\beta_k}}_{\text{own wage}} \times \underbrace{s_{kjt|g}^{\tilde{\sigma}_{kg}}}_{\text{size in market}} \tag{4}$$

with $\tilde{\sigma}_{kg} \equiv 1 - \frac{1}{\sigma_{kg}} \in [0, 1)$, for every $j \in \mathcal{J}_g$.

Appendix F gives the derivation.[15]

The three key parameters. β_k is the wage elasticity of employer choice at the employment margin: the more dispersed idiosyncratic payoffs are, the smaller it is. σ_{kg} applies the same logic within the nest, with $\sigma_{kg} \to 1$ making local employers independent alternatives. The third parameter, ξ_{kjt}, is the unobserved non-wage payoff of firm j to type-k workers—the analogue of

14. The United States, Australia, and New Zealand are grouped with Europe, Mexico with the Americas; workers from the rest of Oceania are excluded.

15. Equation 4 is the quasi-supply function of Berry (1994).

unobserved product quality in demand estimation—and shifts the supply curve firms face, as in Panel (a) of Figure A2.

Firm-specific supply and wage-setting power. Differentiating Equation (4) gives the elasticity of labor supply facing employer j,

$$\mathcal{E}_{kjt} = \beta_k \sigma_{kg} + \beta_k \big[(1 - \sigma_{kg}) s_{kjt|g} - s_{kjt}\big]. \tag{5}$$

If $\mathcal{E}_{kjt} < \infty$, the firm-specific labor supply curve is upward sloping, and employers have wage-setting power, satisfying requirement (i) of Section 4.1. The first term in Equation (5) is the elasticity faced by a small employer ($s_{kjt|g} = s_{kjt} = 0$); the second is negative for $\sigma_{kg} \geq 1$, so the elasticity falls as the firm's share of type-k employment rises, because raising its own wage also raises pay at competing employers. The two terms are the two sources of employer power in requirement (ii) of Section 4.1, as the slope of the firm-specific labor supply curve depends on labor market concentration, both within markets (through $s_{kjt|g}$) and between them (through s_{kjt}), and other factors independent of market structure (through β_k and σ_{kg}).

4.4 Labor demand

The demand side is a wage-posting framework with a finite number of heterogeneous firms. Let l_{kjt} be the labor input of type k at firm j in period t, $l_{.jt} \equiv (l_{1jt}, \ldots, l_{Kjt})$ the vector of inputs, and $C_{jt} \subset \mathcal{K}$ the set of types employed. The technology is

$$F_{jt}(l_{.jt}) = \left(\sum_{k \in C_{jt}} \tilde{\gamma}_{kjt} \, l_{kjt}^{\rho_k} \right)^{\alpha_{jt}}, \tag{6}$$

where $\tilde{\gamma}_{kjt} \equiv \theta_{jt} \gamma_{kjt}$ and $\sum_{k \in C_{jt}} \gamma_{kjt} = 1$. Here $\tilde{\theta}_{jt} \equiv \theta_{jt}^{\alpha_{jt}}$ is total factor productivity (TFP), α_{jt} returns to scale, γ_{kjt} the match-specific productivity of type k at firm j, and ρ_k the curvature of the technology in type-k labor.

This is the productivity heterogeneity satisfying (iii) of Section 4.1. TFP $\tilde{\theta}_{jt}$ varies across firms and time and passes through to wages by way of the marginal revenue product, so the sorting channel of Fact 3 is captured. Match productivity γ_{kjt} differs by firm–type pair, generating Roy sorting (Roy 1951) and giving the imperfect-portability reading of Fact 1 a place in the model. The technology remains additively separable, so $\rho_k < 1$ adds curvature in a type's own input, and this allows for worker types to be imperfect substitutes in production, although the elasticity of substitution is not constant.

Firms post type-specific wages each period, taking their supply curves and other firms' wages as

given.[16] Firm j chooses $\vec{w}_{jt}$ to maximize $P_{jt}F_{jt}(l_{\cdot jt}) - \sum_{k\in\mathcal{K}} w_{kjt}l_{kjt}$, subject to $l_{kjt} = m_{kt}s_{kjt}$ and $\vec{w}_{-j,t}$. The first-order condition rearranges to

$$w_{kjt} = \underbrace{P_{jt}\frac{\partial F_{jt}(l_{\cdot jt})}{\partial l_{kjt}}}_{\text{MRPL}_{kjt}} \times \frac{\mathcal{E}_{kjt}}{1+\mathcal{E}_{kjt}}, \tag{7}$$

and hence to a markdown that depends on the supply elasticity alone:

$$\underbrace{\frac{\text{MRPL}_{kjt} - w_{kjt}}{\text{MRPL}_{kjt}}}_{\text{Wage Markdown}} = \frac{1}{1+\mathcal{E}_{kjt}}. \tag{8}$$

This is the mechanism in Panel (b) of Figure A2. Equation (8) yields a wage and a marginal revenue product for every type at every firm, so the pay gap splits into a gap in marginal revenue products and a gap in the share of that product workers receive. Table A5 collects the notation.

4.5 Identification

I follow the internal instruments strategy of LMS, adapted by CKMM to a setting with heterogeneous worker types and strategic interactions in wage-setting.

4.5.1 The identification problem

The labor supply parameters are identified from

$$\log\frac{s_{kjt}}{s_{k0t}} = \beta_k \log\frac{w_{kjt}}{w_{k0t}} + \tilde{\sigma}_{kg}\log s_{kjt|g} + \xi_{kjt}, \tag{9}$$

which is the log of the quasi-supply function (Lemma 4.1). It holds at any equilibrium, so strategic interactions in wage-setting enter through observables. This matters because firms here are not atomistic: a productivity shock at firm l raises w_{klt}, which induces rival firms j in the same market to raise w_{kjt}. Those general-equilibrium responses change type-k employment at firm j only by changing w_{kjt} and $s_{kjt|g}$, both of which are observed and appear on the right-hand side of Equation (9). Conditioning on $s_{kjt|g}$ is what allows the estimating equation to accommodate strategic interactions, and it is the substantive difference between this specification and the one in LMS.[17]

16. The Bertrand-Nash equilibrium wage vector is unique.

17. The corollary is that a design ignoring $s_{kjt|g}$ and identifying β_k from labor demand shocks alone would require the stable unit treatment value assumption, which strategic interactions violate: firms in the untreated group are themselves affected through their rivals' wage responses. See Berger, Herkenhoff, and Mongey (2022) and CKMM. The approach taken here does not invoke SUTVA.

The main identification challenge is that the supply shifter ξ_{kjt} is unobserved and correlated with both regressors. A firm that is attractive to type-k workers for reasons unrelated to its wage faces a labor supply curve shifted out, as in Panel (a) of Figure A2. It therefore has a larger inside share, a smaller firm-specific elasticity by Equation (5), and posts a lower wage. The direct effect of ξ_{kjt} is thus to attenuate the OLS estimate of β_k.

4.5.2 Which variation identifies which parameter

It is useful to fix ideas with an ideal experiment. Suppose we could shift the productivity of type-k workers at a single firm, $\tilde{\gamma}_{kjt}$, leaving the rest of the market untouched. This shifts firm j's demand for type-k labor and moves the firm along the supply curve it faces, tracing it out. The ratio of the employment response to the wage response recovers $\mathcal{E}_{kjt}$.

What a single such experiment does *not* deliver is a separation of β_k from σ_{kg}, since it recovers only the composite $\mathcal{E}_{kjt}$. Separation requires variation that moves the wage and the inside share separately. Two equivalent ways of seeing this are useful.

First, Equation (5) makes $\mathcal{E}_{kjt}$ a known linear function of the observed shares, with intercept $\beta_k\sigma_{kg}$ and slopes $-\beta_k(\sigma_{kg}-1)$ and $-\beta_k$. Running the experiment at firms of different size within a market therefore identifies $\beta_k\sigma_{kg}$ and β_k separately, and σ_{kg} as their ratio. The intuition is that it is easier for a worker to move to a competing employer within the market than to leave the market or leave employment altogether, so the elasticity facing a small employer exceeds that facing a large one whenever $\sigma_{kg} > 1$, and the gap between them measures how readily type-k workers substitute among local employers.

Second, Equation (9) has two endogenous regressors and therefore requires at least two instruments whose first-stage effects on the pair are not proportional. The production function supplies them, because shocks of different *composition* move the pair differently. An increase in $\tilde{\gamma}_{kjt}$ raises type k's marginal revenue product, so the firm raises w_{kjt} and its type-k employment expands: wage and inside share move together. An increase in $\tilde{\gamma}_{hjt}$ for some other type $h \neq k$ raises the firm's scale, which under decreasing returns *lowers* type k's marginal revenue product, so the pair moves in the opposite direction while the firm's overall employment rises. A shock to firm-wide scale θ_{jt} traces a third ray. Because these shocks displace $(\log w_{kjt}, \log s_{kjt|g})$ along different rays, $(\beta_k, \tilde{\sigma}_{kg})$ are separately identified.

4.5.3 Internal instruments and moment conditions

Making this operational requires a process for productivity. Following CKMM, log labor productivity follows an AR(1) process,

$$\log \tilde{\gamma}_{kjt} = \delta_k \log \tilde{\gamma}_{kjt-1} + \eta_{kjt}, \qquad 0 < \delta_k \leq 1. \tag{10}$$

The process is imposed on $\tilde{\gamma}_{kjt} \equiv \theta_{jt}\gamma_{kjt}$ and not on the normalized term γ_{kjt}.[18] The first assumption is about the first stage of the instrumental variables approach and is testable:

Assumption 4.2 (Relevance)**.** *The variance of type-specific labor productivity shocks is strictly positive,* $\sigma^2_{\eta_k} > 0$ *for each* k*, and* $\delta_k > 0$*.*

The instruments. The instruments are short (one-period) changes in three variables that proxy for the three productivity shocks of Section 4.5.2:

$$\Delta z_{kjt} \equiv z_{kjt} - z_{kjt-1}, \qquad z_{kjt} \in \mathcal{Z} \equiv \Big\{ \log R_{jt},\ \log s_{kjt|g},\ \log s_{\sim kjt|g} \Big\}, \tag{11}$$

where R_{jt} is firm revenue and $\log s_{\sim kjt|g} \equiv \log\big(\sum_{h \in C_{jt},\, h \neq k} s_{hjt|g}\big)$ is the log of the summed inside shares of the firm's other worker types.[19] Revenue growth proxies a shock to firm-wide scale; the own-type inside share proxies a shock to $\tilde{\gamma}_{kjt}$; and the other-types inside share proxies a shock to $\tilde{\gamma}_{hjt}$, $h \neq k$, which shifts the composition of the firm's workforce.

I use internal panel instruments rather than external ones because internal instruments are available for firms in all industries, whereas external instruments are typically available only in narrow contexts, such as the government procurement contracts in Kroft et al. (2024). This matters here because the immigrant–native comparison requires coverage of the full economy rather than of a single sector. Both LMS and CKMM show that internal panel instruments of this form generate results consistent with external instruments used elsewhere in the monopsony literature.

The estimating equation. Writing (9) in long changes gives

$$\Delta_{e,e'} \log \frac{s_{kjt}}{s_{k0t}} = \beta_k \, \Delta_{e,e'} \log \frac{w_{kjt}}{w_{k0t}} + \tilde{\sigma}_{kg} \, \Delta_{e,e'} \log s_{kjt|g} + \Delta_{e,e'} \xi_{kjt}, \qquad \Delta_{e,e'} x_t \equiv x_{t+e} - x_{t-e'}, \tag{12}$$

which is estimated by 2SLS with $e = 2$ and $e' = 3$, instrumenting the two endogenous regressors with Δz_{kjt} for $z \in \mathcal{Z}$. There are three instruments and two parameters per (k, g) cell, so the system is overidentified; the estimation sample conditions on positive type-k employment at both endpoints. Appendix G.1 states the moment conditions.

The asymmetry between the *short* difference of the instrument and the *long* difference of the outcome is the key to the identification strategy. The instrument is dated "tightly": Δz_{kjt} is constructed from variables observed at t and $t-1$ only. The outcome is dated "loosely": its

18. Since $\sum_{k \in C_{jt}} \gamma_{kjt} = 1$ by the normalization in Equation (6), the γ_{kjt} cannot each follow an unrestricted AR(1) process without violating the adding-up constraint. Imposing the process on the unnormalized $\tilde{\gamma}_{kjt}$ is without loss of generality, because the remaining technology parameters are defined residually: $\theta_{jt} = \sum_{k \in C_{jt}} \tilde{\gamma}_{kjt}$ and $\gamma_{kjt} = \tilde{\gamma}_{kjt} / \sum_{h \in C_{jt}} \tilde{\gamma}_{hjt}$. The normalization then holds by construction for any realization of the $\tilde{\gamma}$ process.

19. I use the revenue of the aggregate *Enterprise ID* to construct $\Delta \log R_{jt}$, *not* the imputed revenue of the Enterprise ID–location pair used to identify the production technology parameters in Section 4.5.6.

endpoints sit at $t+2$ and $t-3$, three years on either side of the instrument's window.

4.5.4 The identifying restrictions

Identification requires that the instrument be uncorrelated with the long difference of the shifter. Two conditions on primitives deliver this.

Assumption 4.3 (No anticipation)**.** *Equilibrium revenue, wages and employment shares dated τ are functions of period-τ fundamentals $(\xi_{\cdot\tau},\, \tilde{\gamma}_{\cdot\tau})$ alone.*

Assumption 4.4 (Transitory departures)**.** $\xi_{kjt} = \bar{\xi}_{kj} + u_{kjt} + \psi_k u_{kjt-1}$*, where u_{kjt} is serially uncorrelated.*

Lemma 4.2 (Exogeneity)**.** *Under Assumptions 4.3 and 4.4, for each $z_{kjt} \in \mathcal{Z}$ and for $e \geq 2$, $e' \geq 3$,*

$$\mathrm{Cov}\big(\xi_{kjt+e} - \xi_{kjt-e'},\ \Delta z_{kjt}\big) = 0, \qquad \Delta z_{kjt} \equiv z_{kjt} - z_{kjt-1}.$$

The argument is a statement about dates. Δz_{kjt} is built from equilibrium objects dated t and $t-1$, so by Assumption 4.3 it is a function of u dated t, $t-1$ and $t-2$, while the error is $\Delta_{2,3}\xi_{kjt} = (u_{kjt+2} + \psi_k u_{kjt+1}) - (u_{kjt-3} + \psi_k u_{kjt-4})$. The date sets do not overlap, so the covariance is zero, and the result survives even if u_{kjt} and η_{kjt} are perfectly correlated. Appendix G.2 proves the lemma and shows that $e = 2$, $e' = 3$ is the shortest admissible window.

4.5.5 Threats to identification of the labor supply parameters

What the design does not restrict. There are no restrictions on any firm attributes that are *fixed* over the estimation window. Firms that employ immigrants are not a random sample: they cluster in particular industries and neighborhoods, they are more likely to be co-ethnically owned or managed, to operate in a workplace where a given first language is spoken, or to sit near an established enclave. Each makes a firm persistently more attractive to some type k than to another for reasons unmediated by the posted wage, and each is correlated with the firm's productivity, size, and market. All of these factors are absorbed by $\bar{\xi}_{kj}$ in Assumption 4.4 and differenced out: the identifying restriction concerns the timing of *innovations* to non-wage attractiveness, not their level. What must be ruled out is not that a firm is a persistently good match for a particular group of workers, but that it *becomes* one, permanently, in the same years its productivity changes—which is the network story discussed next.

Networks and permanence of the supply shifter. The main threat is related to networks and a potential violation of Assumption 4.4. Suppose a firm that expands after a productivity shock hires immigrant workers, and that those workers refer co-nationals to the firm in later

years. The firm's non-wage standing with type k then rises *durably* in response to the shock: ξ_{kjt} contains a component with more persistence than MA(1) and $\text{Cov}(\Delta_{e,e'}\xi_{kjt}, \Delta z_{kjt}) > 0$.

However, the direction of the resulting bias is unambiguous and works *against* the paper's main findings. A positive covariance loads employment growth that is in truth a supply-shifter response onto the wage and inside-share terms of Equation (9), biasing $\hat{\beta}_k$ and $\hat{\tilde{\sigma}}_{kg}$ *upward*. By Equation (5) the firm-specific elasticity $\mathcal{E}_{kjt}$ is then overstated and, by Equation (8), the markdown is *understated*. Referral networks of this kind are widely documented to be stronger among immigrants than among natives (Munshi 2003; Dustmann et al. 2016), so the bias should be larger for immigrant types. The estimated immigrant–native gap in markdowns is therefore a lower bound on the true gap, and the share of the immigrant–native pay gap attributed to labor market power in Section 5 is conservative.

Anticipation and the timing of hiring. Assumption 4.3 would fail if firms acted on productivity innovations before they occurred, or if their hiring of type k workers drifted in anticipation of a productivity shock. What is at stake is a window: the instrument is dated t and $t-1$, while the long difference reaches back to $t-3$, so the assumption requires only that the instrument carry no information about anything dated $t-3$ or earlier. The test is therefore whether treated and control firms had already begun to separate before the base period, and Appendix G.3 shows they had not, in an exercise analogous to Figure 1 of LMS: splitting firms on revenue growth and on revenue per type-k worker, the two groups are indistinguishable in the years running up to $t-3$, diverge sharply at the event, and run parallel after $t+2$.

4.5.6 Labor demand parameters

Given the labor supply parameters, the labor supply elasticity $\mathcal{E}_{kjt}$ follows from Equation (5), and the firm's first-order condition (7) then delivers the marginal revenue product of labor:

$$\text{MRPL}_{kjt} = \frac{1 + \mathcal{E}_{kjt}}{\mathcal{E}_{kjt}} w_{kjt}. \tag{13}$$

Equation (13) returns a marginal revenue product for every firm–type–year cell directly from the posted wage and the estimated elasticity. The next step is to identify the production technology parameters, given Equation (13).

Substitutability across types. The first technology parameter is the curvature ρ_k. Differentiating Equation (6) and taking the ratio of the first-order conditions for two types $k, h \in C_{jt}$

removes the firm's scale, its TFP and its returns to scale in one step:

$$\frac{\text{MRPL}_{kjt}}{\text{MRPL}_{hjt}} = \frac{\tilde{\gamma}_{kjt}\,\rho_k\,l_{kjt}^{\rho_k-1}}{\tilde{\gamma}_{hjt}\,\rho_h\,l_{hjt}^{\rho_h-1}}. \tag{14}$$

Taking logs gives a log-linear relationship in which each type's curvature is the coefficient on its own employment:

$$y_{khjt} \;\equiv\; \log\frac{\text{MRPL}_{kjt}}{\text{MRPL}_{hjt}} = \log\frac{\rho_k}{\rho_h} + (\rho_k-1)\log l_{kjt} - (\rho_h-1)\log l_{hjt} + \omega_{khjt}, \tag{15}$$

with $\omega_{khjt} \equiv \log(\tilde{\gamma}_{kjt}/\tilde{\gamma}_{hjt})$ the unobserved relative productivity of the two types at the firm. That unobservable is precisely what makes the firm hire the quantities on the right-hand side, so OLS estimation of Equation (15) is biased.

The productivity process already imposed in Equation (10) presents a solution, subject to one restriction: I follow CKMM in setting $\delta_k = \delta$ for all k, since type-specific persistence is not separately identified from ρ_k here.[20] Applying Equation (10), it is possible to write the residual as $\omega_{khjt} = \delta\,\omega_{khjt-1} + \zeta_{khjt}$, where $\zeta_{khjt} \equiv \eta_{kjt} - \eta_{hjt}$ is a difference of innovations dated t.

Subtracting δ times the period-$(t-1)$ version of Equation (15) from the period-t version then removes ω_{khjt-1} and leaves the innovation as the only unobservable:

$$y_{khjt} - \delta\,y_{khjt-1} = c_{kht} + (\rho_k-1)\big(\log l_{kjt} - \delta\log l_{kjt-1}\big) - (\rho_h-1)\big(\log l_{hjt} - \delta\log l_{hjt-1}\big) + \zeta_{khjt}. \tag{16}$$

The intercept c_{kht} absorbs both the pair-specific constant $(1-\delta)\log(\rho_k/\rho_h)$ and any aggregate movement in the relative productivity of types k and h, and is implemented as a pair-by-year fixed effect.

Because ζ_{khjt} is dated t and Equation (10) implies innovations are serially uncorrelated, anything dated $t-1$ or earlier is orthogonal to it. The first lags of employment and of the MRPL ratio already appear on the right-hand side of Equation (16) and are predetermined; their squares and cross-products therefore serve as excluded instruments.

Equation (16) holds for every pair (k,h), and the same ρ_k appears in every equation involving type k. Rather than impose the nonlinearity in δ directly, I follow CKMM in estimating the compound parameters $\rho_k - 1$ and $\delta(\rho_k - 1)$ as free coefficients, which makes the system linear in $2K+1$ parameters, and stacking the resulting moments across all pairs, firms and years. Estimation is by 2SLS, and ρ_k is recovered post-estimation. Because the restriction linking the two compound parameters is not imposed, each ρ_k can be computed twice, which delivers a

20. The restriction is needed only for identifying the curvature of the production technology; the labor-supply moments of Section 4.5.3 do not require restrictions on δ_k.

specification test of the productivity process at no cost (see Appendix G.4 for details). Perfect substitutability is the joint null $\rho_k = 1$ for all k, and Section 5 rejects it for every type.

Match productivity, returns to scale and TFP. The remaining parameters follow from $\hat{\rho}_k$ by accounting, with no further assumptions beyond a common output price. Define $z_{kjt} \equiv \text{MRPL}_{kjt}/(\rho_k\, l_{kjt}^{\rho_k - 1})$ which strips from each type's MRPL the part that reflects how much of that type the firm has already hired. Equation (14) says $z_{kjt}/z_{hjt} = \gamma_{kjt}/\gamma_{hjt}$, so the normalization $\sum_{k \in C_{jt}} \gamma_{kjt} = 1$ pins down the level: $\gamma_{kjt} = z_{kjt}/(\sum_{h \in C_{jt}} z_{hjt})$. Writing $\Lambda_{jt} \equiv \sum_{h \in C_{jt}} \gamma_{hjt} l_{hjt}^{\rho_h}$ for the effective labor aggregate, returns to scale and TFP then follow from the marginal product and the production function respectively: $\alpha_{jt} = (\Lambda_{jt} \sum_h \text{MRPL}_{hjt}\, l_{hjt})/(R_{jt} \sum_h \rho_h\, \gamma_{hjt}\, l_{hjt}^{\rho_h})$ and $\tilde{\theta}_{jt} = R_{jt}/\Lambda_{jt}^{\alpha_{jt}}$. Revenue here is that of the local Enterprise ID–location unit, since each local unit is allowed its own technology.[21]

4.6 Estimation

Because Equation (12) is written in logs of firm–type employment shares, it requires strictly positive type-k employment at firm j at both endpoints of the long difference, and firm–type–year cells that fail this restriction are dropped. However, this restriction governs *estimation* only. The estimated parameters $(\hat{\beta}_k, \hat{\tilde{\sigma}}_{kg}, \hat{\rho}_k)$ vary at the type and market level, so once they are in hand I map them back onto the *full* AKM sample, calculate the remaining productivity parameters $(\alpha_{jt}, \theta_{jt}, \gamma_{kjt})$, and evaluate average markdowns and any other differences between immigrants and natives there. Every number reported in Section 5 therefore describes the full AKM sample, not the subset on which estimation is feasible. That mapping is valid provided the estimation subset does not select on the parameters themselves.

Table A6 reports summary statistics for the estimation sample for comparison with the AKM sample in Table 1. The key gaps in earnings and productivity are very similar across the two samples: the immigrant–native gap in log revenue per worker is -0.05 in the AKM sample and -0.06 in the estimation sample, and the gap in log mean earnings is -0.184 and -0.197.

Three other patterns stand out. First, retention is high, and higher for immigrants than for natives: 75.8 percent of person-years survive, 74.4 percent of native and 81.5 percent of permanent-resident person-years. Second, composition is preserved. The Economic, Family and Refugee shares of permanent-resident person-years move by less than a tenth of a percentage point; gender composition moves by at most 1.3 percentage points and the provincial distribution by at most 0.7. Third, the samples differ in employer size: the requirement of positive type-k employment at both endpoints binds hardest at small firms, so median employer size rises from 29 to

21. This is in contrast to the revenue of the *aggregate* Enterprise ID used to construct the supply-side instruments in Equation (11).

47 workers for natives and from 75 to 96 for permanent residents.

However, the shift to slightly larger firms is not a threat to the mapping, because firm size enters the model as an observable rather than as an unmodeled feature of the sample. What is estimated is the pair (β_k, σ_{kg}), which is common within a type–market cell; size affects the firm-specific elasticity only through the inside shares, which Equation (5) makes an explicit argument of $\mathcal{E}_{kjt}$ and which Equation (12) conditions on directly. When the estimated parameters are mapped back to the full AKM sample, each firm–type–year cell is evaluated at its own observed shares, so small employers receive the elasticities and markdowns their own shares imply.

Four other choices are worth discussing. First, the nesting parameter σ_{kg} cannot be estimated for every type–market pair for computational reasons. The heterogeneity in σ_{kg} is restricted to four province groups: Ontario, Quebec, British Columbia, and the rest of Canada, so that conditional on type k, every local market g in a province group has the same σ_{kg}.[22] Second, for all parameters that vary across firms, I report employment-weighted averages. For example, the "average immigrant markdown" refers to the markdown faced by the average immigrant *worker* rather than by the average firm employing immigrants in the AKM sample. I do this because the aggregate pay gap is itself worker-weighted. Third, the confidence intervals reported are at the 95 percent level and are obtained using the Hall (1992) bootstrap procedure, and these intervals may be asymmetric if the underlying distribution is skewed (Appendix H). Finally, earnings and revenue are winsorized at the 0.5 and 99.5 percentiles before estimation.

5 Results

5.1 Labor Supply and the Wage Markdown

Monopsony power is pervasive. By Equation (8) the markdown is a function of the firm-specific labor supply elasticity, so I begin there; Appendix I discusses the results for the model primitives. Table 4 shows that the employment-weighted average elasticity across all workers is 5.25, with a 95 percent bootstrap confidence interval from 5.09 to 5.38: an employer that cuts pay by one percent loses roughly five percent of its workforce. The estimate lies within the range of credible firm-level estimates, which typically fall between three and six (Card 2022; Sokolova and Sorensen 2021; Manning 2021). Estimating Equation (4) by OLS instead yields 0.289. The bias has the anticipated sign: firms offering more attractive non-wage amenities attract more workers and can therefore pay less, so the OLS slope understates the elasticity.

22. I choose province groups for the main specification because geographic clustering is an important factor for immigrant earnings (see Fact 4 in Section 3).

Immigrant labor supply is more inelastic. Employers of immigrant labor face an average elasticity of 4.42 against 5.45 for native labor, with confidence intervals from 4.06 to 4.72 and from 5.29 to 5.60 respectively (Panel B of Table 4).[23] An employer that lowers wages therefore retains a larger share of its immigrant workforce than of its native workforce, the configuration drawn in Figure A2b. Panel C disaggregates the immigrant sample: the ordering is 5.09 for economic-class immigrants, 3.85 for family-class immigrants, and 3.20 for refugees. The interval for the economic class overlaps that for natives but not those for the other two categories, which overlap each other. The ordering reproduces the ranking of the conditional pay gaps in Fact 1 and tracks the degree to which admission is conditioned on labor market attachment.

Two sources of wage-setting power. Equation (5) shows that the firm-specific elasticity is also a function of employer size, since a large employer that raises its wage also raises pay at the alternatives its workers would move to. Figure 4 shows that elasticities fall, and markdowns rise, with the employer's share of its local labor market, for both groups. Immigrants' elasticities lie below natives' at every share and their markdowns above. The shared slope is wage-setting power operating through market structure; the vertical gap is wage-setting power from everything else. That gap is not constant: the immigrant–native difference in markdowns widens from 7.3 percentage points among employers with the smallest local shares to 12.2 points among those with the largest. Market structure and substitutability are therefore not additive, a point I return to in Section 5.4.

Markdowns, and the structural counterpart to Fact 5. Equation (8) converts these elasticities into the share of the marginal revenue product withheld. Table 4 shows that employers in Canada withhold 17.6 percent on average, but 16.4 percent from natives against 22.8 percent from immigrants, with intervals of 16.1 to 16.8 and 21.6 to 23.4, respectively. The resulting 6.4 percentage point gap is close to the differential of roughly 7.7 log points that Hirsch and Jahn (2015) estimate for Germany. Moreover, the 6.4 percentage point gap in markdowns is consistent with Fact 5, which showed that immigrant-intensive firms pay roughly 6.8 log points less at every level of revenue per worker.[24] By admission category the wage markdown estimates are 20.6, 24.6, and 27.0 percent for economic-class immigrants, family-class immigrants, and refugees.

23. Native-born men face an average elasticity of 5.63 against 5.09 for native-born women, with 95 percent confidence intervals that do not overlap, consistent with Sharma (2024), Webber (2016), and the conjecture of Robinson (1933).

24. The two magnitudes are not directly comparable. Fact 5 conditions on revenue per worker, an average product, whereas the markdown in Equation (8) is defined against the marginal revenue product. The two coincide only under constant returns, and Section 5.2 shows $\hat{\alpha}_{jt} < 1$. Fact 5 also compares firms rather than workers, and Appendix D.3 shows that the component of the markdown common to all of a worker's potential employers does not enter the AKM firm effect. The agreement in magnitude is therefore a coincidence of two distinct comparisons, and I treat Fact 5 as corroborating the sign and rough order of magnitude rather than as a validation of the estimate itself.

5.2 The Production Technology

The markdown estimates described in Section 5.1 imply a gap in marginal revenue products of 5.3 percent, against a pay gap of 15.0 percent. The remainder of this section explores the estimates of the underlying technology parameters and the immigrant–native differences in each.

Returns to scale and the comparison to Fact 3. Table 4 shows that average returns to scale are 0.291, with a confidence interval from 0.289 to 0.292, substantially decreasing and in line with CKMM and LMS. The immigrant and native averages, 0.295 and 0.290, have confidence intervals that do not overlap, but the difference is small: economically, immigrants and natives work at employers whose technologies exhibit the same curvature in total labor input. Among admission categories, employers of economic-class immigrants have somewhat higher returns to scale than the rest, while family-class immigrants and refugees are statistically indistinguishable from one another. Furthermore, Fact 3 compared revenue per worker, an average product, while wage setting responds to the marginal product, and with α_{jt} well below one the two differ by a factor that is large and varies across employers. Fact 3 is therefore suggestive of a productivity gap but does not measure the object entering Equation (7).

Immigrant and native substitutability. The average ρ_k is 0.888, with a confidence interval from 0.887 to 0.890 (Table 4), so worker types are imperfect substitutes and each type's marginal product declines in its own employment. The estimate for immigrant labor, 0.930, exceeds that for native labor, 0.878, with economic-class immigrants at 0.916, refugees at 0.941, and family-class immigrants at 0.947. Every estimate is bounded away from one at the 95 percent level, and the intervals are mutually non-overlapping across the three admission categories as well as between immigrants and natives.[25]

The levels of $\tilde{\theta}_{jt}$ and γ_{kjt}. Average $\log \tilde{\theta}_{jt}$ is 16.5 for immigrants against 15.8 for natives, and average γ_{kjt} is 0.172 against 0.424 (Appendix Table A15). Both differences are precisely estimated, but neither should be read as a productivity ranking. The two parameters enter Equation (6) only through the product $\tilde{\theta}_{jt}\gamma_{kjt}$, and are separated by the adding-up restriction $\sum_{k \in C_{jt}} \gamma_{kjt} = 1$. An employer drawing on a wider set of types records a smaller γ_{kjt} on each and, holding output fixed, a larger $\tilde{\theta}_{jt}$. The between-firm sorting documented in Fact 2 is relevant here: immigrants work at employers that are larger and draw on more types (Table 1), so a high $\tilde{\theta}$ and a low γ reflect sorting patterns and segregation.

25. The magnitudes are modest, but the ordering contradicts an alternative explanation for the pay gap. Were immigrant labor to encounter diminishing returns more rapidly, the employers that concentrate immigrants—the low-premium firms of Fact 2—would face lower marginal products and would pay less for a purely technological reason.

5.3 Exposure to Labor Market Concentration

Fact 4 measured exposure to concentration with the HHI and the collision probability, which are functions of employer shares. Intuitively, labor concentration matters for a worker because it proxies for the scarcity of that worker's outside options. Yet, an employer is an outside option only if the worker would actually work there. A market split evenly among many employers may look unconcentrated according to the HHI or collision probability, even if only a few firms are actual employment possibilities for an individual worker. The generalized concentration index (GCI) of CKMM weights employers by how substitutable they are for the worker, and therefore it is a broad summary of exposure to wage-setting power. For worker type k in period t:

$$GCI_{kt} \equiv \underbrace{\prod_{g \in \mathcal{G}} \left(\exp \left\{ \sum_{j \in \mathcal{J}_g} s_{kjt|g} \log s_{kjt|g} \right\} \right)^{\frac{s_{kgt}}{\sigma_{kg}}}}_{\text{Within-market index (WGCI)}} \times \underbrace{\exp \left\{ \sum_{g \in \mathcal{G}} s_{kgt} \log s_{kgt} \right\}}_{\text{Between-market index (BGCI)}}. \tag{17}$$

Expected worker welfare is decreasing in GCI_{kt}, so a gap between two groups is a statement about what market structure costs them (Appendix I.4.1).

Immigrants' overall exposure is 9.14 (SE = 0.03) against 4.90 (SE = 0.01) for natives, a ratio of 1.87 (Table 5, all indices scaled by 100). The composition follows Fact 4: the between-market component is exactly twice as high for immigrants, 11.46 against 5.73, while the within-market component is lower, 77.04 against 84.98.[26] All three admission categories are more exposed than the native-born, with refugees highest at 9.81, followed by family-class immigrants at 9.58 and economic-class immigrants at 8.69, reproducing the ordering of the markdowns in Table 4.

Immigrants' disadvantage is larger than the descriptive statistics suggested: 1.87 times the native level using the GCI compared to 1.33 times using the collision probability. The entire discrepancy is within-market. Using employer shares, immigrants seemed to be less exposed to within-market concentration because they work in larger cities; weighting those employers by how substitutable they are to an individual immigrant worker, however, the advantage mostly disappears.

5.4 Discussion

What the elasticities reflect. How do we interpret the lower firm-specific elasticities for immigrants compared to natives, driven by concentration across employers but also the labor supply primitives β_k and σ_{kg}?

Equation 5 shows that the product $\beta_k \sigma_{kg}$ enters directly into the elasticity formula. The first

26. The two components do not multiply to the overall index in Table 5 because each is an employment-weighted average of a group-year index; within each worker type and year the decomposition is exact.

interpretation of $\beta_k\sigma_{kg}$ follows Assumption 4.1 literally. A low $\beta_k\sigma_{kg}$ reflects dispersed idiosyncratic tastes, so that local employers are poor substitutes for one another in the eyes of a type-k worker. This need not be a statement about preferences in a narrow sense. It says that, whatever the number of employers in the market, few of them are suitable alternatives for the worker. In this sense, a low $\beta_k\sigma_{kg}$ means that the worker has few outside options.

The second interpretation is informational. Appendix F.2 shows that a model without taste shocks, in which workers know every local employer but must pay to learn what each one offers, yields Equation (4) with the same functional form and the same parameters. In that model, β_k indexes how costly it is to learn about pay, σ_{kg} indexes how costly it is to compare employers within a market, and ξ_{kjt} measures how visible firm j is to type-k workers. The equivalence is not surprising. A choice system reveals how substitutable the alternatives are, not why they are substitutable (Manning 2003; Card et al. 2018).

The informational interpretation may be especially natural in this setting. Immigrants may find it more difficult to learn about the host-country labor market. They typically know fewer local employers, have less information about what those employers pay, may search in a second language, and often learn about openings through co-national networks that cover a narrow set of firms (Munshi 2003; Dustmann et al. 2016). Each of these features makes it harder to compare employers and lowers $\beta_k\sigma_{kg}$.

The lower estimates of $\beta_k\sigma_{kg}$ for immigrants should therefore be read as a summary of several mechanisms, including both preferences and information, that limit their outside options and mobility across employers. The results in Section 5 do not require separating these mechanisms. What they establish is that, for immigrant workers, other local employers are poorer substitutes for their current employer than they are for comparable native workers. As a result, employers face a steeper immigrant supply curve and mark down immigrant wages more.

Concentration, substitutability, and welfare. The markdown in Equation (8) depends on the residual elasticity $\mathcal{E}_{kjt}$, which combines σ_{kg} with the employer's share of type-k employment in the market. Substitutability and concentration are therefore not competing explanations of wage-setting power but the two arguments of the same function: a low σ_{kg} makes a given share more valuable to the firm, and a high share makes a given σ_{kg} more binding.

6 Decomposing the Pay Gap

Section 5 decomposed the pay gap into a markdown differential and a differential in marginal revenue products. That decomposition is an identity: it holds at the estimated parameters and the observed allocation of workers across firms, and it apportions the gap between the two terms.

However, the markdown differential is not the amount by which the gap would fall if employers lost their wage-setting power over immigrants, because removing it would change where immigrants work and therefore the marginal products entering the second term.

This section decomposes the pay gap using a general equilibrium framework. Each experiment either replaces a set of parameters with a common value or restricts the firm's wage-setting rule, and the model is re-solved for the equilibrium wage vector. Workers reallocate across employers, labor market shares update, and markdowns and marginal revenue products adjust jointly. Shapley values over the resulting set of experiments then attribute the gap across channels. The four channels are chosen so that each corresponds to a distinct policy lever, and the decomposition therefore says how much of the gap each lever can influence.

6.1 Four channels

Each counterfactual shuts down one channel through which immigrants and natives can end up with different wages, holding the rest of the model at its estimated values. Three of the four do so by replacing a set of parameters with a common value; the fourth changes the wage-setting rule itself.

- *Employer substitutability* ($\mathcal{S}$): β_k and σ_{kg}, which govern the extent to which employers are substitutable and therefore the slope of the supply curve. Equalizing them removes the differential in wage-setting power that operates independently of market structure.
- *Market structure* ($\mathcal{M}$): the strategic term in Equation (5). Rather than equalize a parameter, this experiment restricts the firm's problem, setting employment shares to zero in the perceived elasticity so that each employer behaves atomistically. It removes the differential that operates through market structure.
- *Value of non-employment* ($\mathcal{V}$): w_{0k}, the parameter related to the social safety net.
- *Skill mix and technology* ($\mathcal{R}$): the type populations m_k, the supply shifters ξ_{kj}, the match productivities γ_{kj}, and the technology parameters $\tilde{\theta}_j$, α_j and ρ_k. These determine each firm's workforce and what it produces. Workers sort across firms on both ξ_{kj} and γ_{kj}, so the two are restricted together.[27] These parameters change markdowns only by changing employment shares, the channel $\mathcal{M}$ isolates. Equalizing them removes the differential in marginal revenue products that is not due to sorting on wage-setting power.

Each block corresponds to a distinct policy margin, which is what makes the decomposition informative. $\mathcal{S}$ collects the frictions that make one employer a poor substitute for another at a given level of concentration, such as information asymmetries that disadvantage immigrants

27. They are also not separately identified when $s_{kj} = 0$: a zero share of type k at firm j is consistent with low attachment ($\xi_{kj} = -\infty$) or with low productivity ($\gamma_{kj} = 0$).

relative to natives (see Section 5.4); the policies that act on it are those that improve information and mobility. $\mathcal{M}$ is market structure, the object of antitrust enforcement, merger review, and restrictions on non-compete agreements. $\mathcal{V}$ is the value of non-employment, which depends on settlement supports and other policies related to the social safety net. $\mathcal{R}$ is determined by immigrant selection criteria, training, and industrial policy.

6.2 Design

I restrict the data to 2016 for computational feasibility.[28] Each counterfactual replaces a subset of parameters with a common value or alters the wage-setting rule and solves for the fixed point of Equations (4) and (7) using the under-relaxed Jacobi iteration of Appendix J.1.[29] Equalizing $\mathcal{R}$ is thus the experiment in which immigrants arrive with the skills of natives and can deploy them at the same employers, which is the limit of what selection and credential recognition policy can deliver. Wages, markdowns, marginal products and the allocation of workers all adjust. The outcome is the employment-weighted gap

$$\mathcal{D} \;=\; 1 - \frac{\sum_{k\in\mathcal{I}}\sum_j l_{kj}\, w_{kj} \big/ \sum_{k\in\mathcal{I}}\sum_j l_{kj}}{\sum_{k\in\mathcal{N}}\sum_j l_{kj}\, w_{kj} \big/ \sum_{k\in\mathcal{N}}\sum_j l_{kj}}, \tag{18}$$

where $\mathcal{I}$ and $\mathcal{N}$ index immigrant and native types and w and l take their counterfactual values. Appendix Table A7 summarizes the four blocks. The observed gap in the data for 2016 is 14.99%, and imposing all four restrictions at once moves the gap to -0.27%. Appendix J.2 reports the solver's reproduction of the observed gap at the estimated parameters.

6.3 Results

The blocks interact, so each block's contribution depends on the order of elimination. I therefore report each block's Shapley value, which averages its marginal contribution over all $4! = 24$ elimination orders,

$$\phi_b \;=\; \sum_{\mathcal{C}\subseteq\mathcal{B}\setminus\{b\}} \frac{|\mathcal{C}|!\,\big(|\mathcal{B}| - |\mathcal{C}| - 1\big)!}{|\mathcal{B}|!}\Big[\mathcal{D}(\mathcal{C}\cup\{b\}) - \mathcal{D}(\mathcal{C})\Big], \tag{19}$$

28. The fixed point is solved once per counterfactual, for fifteen counterfactuals; a full panel is infeasible. 2016 sits mid-sample and is the year of the Census evidence in Section 3.

29. For parameters defined at the type level, such as β_k, the common value is the value for male natives. For parameters that vary across firms rather than individuals, such as α_j and $\tilde{\theta}_j$, I take the employment-weighted median among male natives and assign it to all firms. For the match terms γ_{kj} and ξ_{kj}, the common value is that of male natives at the same firm, so that an immigrant type is as productive at employer j, and as willing to work there, as a native is. For counterfactual $\mathcal{M}$, I shut down strategic interactions by assuming that the elasticity does not depend on shares.

where $\mathcal{B} = \{\mathcal{S}, \mathcal{M}, \mathcal{V}, \mathcal{R}\}$ and $\mathcal{D}(\mathcal{C})$ is the counterfactual gap when the blocks in $\mathcal{C}$ are shut down. Averaging over orders is necessary because shutting a block down first versus shutting it down last generally produces different numbers, so any one order makes each block's contribution an artifact of where it happened to fall in the sequence. Shutting each block down on its own, from the baseline in every case, avoids that but leaves a residual the four blocks do not account for and therefore does not produce a proper decomposition of the pay gap. The Shapley average is the only rule that avoids both problems (Shorrocks 2013), and its four contributions sum to the -15.26 pp that take the gap from 14.99% to essentially zero.

Figure 5a plots each Shapley value together with the range of the eight marginal contributions.[30] Employer substitutability accounts for 74.6% of the gap, market structure for 6.3%, the value of non-employment for 7.9%, and skill mix and technology for 11.3%. Wage-setting power is the sum of the first two: it accounts for 80.8% overall. The policy reading is that selection and industrial policy, the levers that act on $\mathcal{R}$, influence at most about a ninth of the gap, and settlement supports a comparable share, while the bulk of the gap is reached only by policies that change how employers compete for immigrant workers.

Figure 5b reports a further pattern. Every counterfactual that equalizes employer substitutability reduces the gap in marginal products, in most cases reversing its sign, even when no technology parameters are altered. This happens through reallocation. In the baseline, less productive firms post lower wages, and since immigrants have lower elasticities, they end up disproportionately employed in low productivity matches. Eliminating employer wage-setting power is therefore not purely redistributive: immigrants both retain a larger share of their marginal revenue product *and* reallocate toward more productive employers. This is a striking result that demonstrates the importance of decomposing the pay gap in general equilibrium. It also means that the gap in marginal products is not a fixed feature of the immigrant workforce to be addressed by selection policy alone: part of it is produced by the wage-setting environment itself.

Strategic interactions appear to contribute little on their own, but that understates their role. Eliminating them from the baseline closes 0.66 pp of the gap. Eliminating them after employer substitutability has been equalized closes 4.29 pp, six times as much. The reason is that the first experiment changes where immigrants work. Once employers are equally substitutable, immigrants sort toward more productive firms, and those firms hold larger shares of the markets they hire from. Immigrants' exposure to concentration therefore rises, and market structure becomes a larger part of what remains of the gap. With strategic interactions in play, part of what equalizing employer substitutability accomplishes is thus offset. Once firm productivity is equalized as well and there are no more productive employers to sort toward, eliminating

30. Appendix J discusses the full matrix of 32 marginal contributions averaged to produce the Shapley decomposition in Figure 5.

strategic interactions closes only 0.02 pp. Antitrust enforcement and mobility policy are therefore complements: the return to the first rises once the second is in place.

7 Conclusion

This paper asks how much of the immigrant–native pay gap in Canada is generated by employer wage-setting power rather than by differences in productivity. Several facts frame the analysis: the gap widens rather than closes conditional on observed skill, immigrants are more exposed to employer concentration, and immigrant-intensive employers pay less than comparably productive native-intensive firms at every point of the productivity distribution. An oligopsony model estimated on the universe of Canadian tax filers shows that firms withhold 22.8 percent of the marginal revenue product from immigrants against 16.4 percent from natives, with the differential ordered across admission classes in the same way as the conditional pay gaps. Marginal products differ by only 5.3 percent. A general equilibrium decomposition attributes 80.8 percent of the gap to employer wage-setting power and shows that wage-setting power leads to misallocation that concentrates immigrants in low productivity firms.

Three implications follow. First, the gap is not primarily a productivity gap, so policies to reduce the pay gap should focus on eliminating frictions that generate wage-setting power: marginal products differ by 5.3 percent while pay differs by 15.0. The estimates point to the set of frictions that make one employer a poor substitute for another, such as language barriers or informational asymmetries, which account for three quarters of the pay gap in the decomposition. Second, the lever Canadian policy has focused on most—immigrant selection criteria—is not sufficient to close the immigrant-native pay gap on its own. Skill mix and technology accounts for 11.3 percent of the gap, and the value of non-employment for a further 7.9 percent. Selecting more educated immigrants raises what they can produce, but leaves untouched the share of it they are paid. Third, because wage-setting power also concentrates immigrants in lower productivity matches, reducing it will not result in a pure transfer from firms to workers. Equalizing employer substitutability reverses the sign of the marginal product gap without altering a single technology parameter, so the same policies that reduce wage-setting power will raise aggregate output as well as immigrant pay. Which of the specific frictions matters most, and what reduces them at acceptable cost, are questions the decomposition poses but does not answer. Investigating this further is the natural next step for research on the economic integration of immigrants.

References

Abowd, JM, F Kramarz, and DN Margolis. 1999. "High wage workers and high wage firms." *Econometrica* 67 (2): 251–333.

Albert, Christoph, and Joan Monras. 2022. "Immigration and Spatial Equilibrium: The Role of Expenditures in the Country of Origin." *American Economic Review* 112 (11): 3763–3802.

Amior, Michael, and Alan Manning. 2020. "Monopsony and the wage effects of migration." *CEP Discussion Papers.*

Amior, Michael, and Jan Stuhler. 2024. "Immigration, Monopsony and the Distribution of Firm Pay."

Amo-Agyei, Silas. 2020. "The migrant pay gap: Understanding wage differences between migrants and nationals," https://www.ilo.org/sites/default/files/wcmsp5/groups/public/@ed_protect/@protrav/@migrant/documents/publication/wcms_763803.pdf.

Arellano-Bover, Jaime, and Shmuel San. 2024. "The Role of Firms and Job Mobility in the Assimilation of Immigrants: Former Soviet Union Jews in Israel."

Åslund, Olof, Cristina Bratu, Stefano Lombardi, and Anna Thoresson. 2025. *Firm Productivity, Manager Origin, and Immigrant-Native Earnings Disparities.* Working Paper SSRN Working Paper 5349339. Social Science Research Network. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5349339.

Beach, Charles M., Alan G. Green, and Christopher Worswick. 2011. *Toward Improving Canada's Skilled Immigration Policy: An Evaluation Approach.* Policy Study 45. Toronto: C.D. Howe Institute.

Beauregard, Pierre-Loup, Thomas Lemieux, Derek Messacar, and Raffaele Saggio. 2025. *Why Do Union Jobs Pay More? New Evidence from Matched Employer-Employee Data.* NBER Working Paper 33740. Conditionally accepted, *Quarterly Journal of Economics.* National Bureau of Economic Research.

Berger, David, Kyle Herkenhoff, and Simon Mongey. 2022. "Labor Market Power." *American Economic Review* 112 (4): 1147–1193.

Berry, Steven T. 1994. "Estimating Discrete-Choice Models of Product Differentiation." *The RAND Journal of Economics* 25 (2): 242–262.

Blinder, Alan S. 1973. "Wage Discrimination: Reduced Form and Structural Estimates." *The Journal of Human Resources* 8 (4): 436–455.

Card, David. 2001. "Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration." *Journal of Labor Economics* 19 (1): 22–64.

———. 2022. "Who Set Your Wage?" *American Economic Review* 112 (4): 1075–1090.

Card, David, Ana Rute Cardoso, Joerg Heining, and Patrick Kline. 2018. "Firms and Labor Market Inequality: Evidence and Some Theory." *Journal of Labor Economics* 36 (1): S13–S70.

Card, David, Ana Rute Cardoso, and Patrick 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.

Chan, Mons, Kory Kroft, Elena Mattana, and Ismael Mourifie. 2024. "An Empirical Framework For Maching with Imperfect Competition."

Chiswick, B. R., and P. W. Miller. 1995. "The Endogeneity Between Language and Earnings - International Analyses." *Journal of Labor Economics* 13 (2): 246–288.

Damas de Matos, Ana. 2017. "Firm heterogeneity and immigrant wage assimilation." *Applied Economics Letters* 24 (9): 653–657.

Damm, Anna Piil. 2009. "Ethnic Enclaves and Immigrant Labor Market Outcomes: Quasi-Experimental Evidence." *Journal of Labor Economics* 27 (2): 281–314.

Depew, B, P Norlander, and TA Sorensen. 2017. "Inter-firm mobility and return migration patterns of skilled guest workers." *Journal of Population Economics* 30 (2): 681–721.

Dostie, Benoit, Jiang Li, David Card, and Daniel Parent. 2023. "Employer policies and the immigrantânative earnings gap." *Journal of Econometrics* 233 (2): 544–567.

Dustmann, Christian, Albrecht Glitz, Uta Schönberg, and Herbert Brücker. 2016. "Referral-based Job Search Networks." *The Review of Economic Studies* 83 (2): 514–546.

Edin, Per-Anders, Peter Fredriksson, and Olof Åslund. 2003. "Ethnic Enclaves and the Economic Success of Immigrants—Evidence from a Natural Experiment." *The Quarterly Journal of Economics* 118 (1): 329–357.

Ferrer, A, DA Green, and WC Riddell. 2006. "The effect of literacy on immigrant earnings." *Journal of human resources* 41 (2): 380–410.

Fortin, Nicole, Thomas Lemieux, and Javier Torres. 2016. "Foreign human capital and the earnings gap between immigrants and Canadian-born workers." *Labour Economics* 41:104–119.

Gerard, FranÃ§ois, Lorenzo Lagos, Edson Severnini, and David Card. 2021. "Assortative Matching or Exclusionary Hiring? The Impact of Employment and Pay Policies on Racial Wage Differences in Brazil." *American Economic Review* 111 (10): 3418–3457. Accessed August 25, 2024. https://www.aeaweb.org/articles?id=10.1257/aer.20181596.

Gyetvay, Sam, and Sekou Keita. 2024. "Competition, Mobility and Immigration."

Hall, Peter. 1992. *The Bootstrap and Edgeworth Expansion.* Springer Series in Statistics. New York, NY: Springer.

Hirsch, B, and EJ Jahn. 2015. "Is There Monopsonistic Discrimination Against Immigrants?" *ILR Review* 68 (3): 501–528.

Javdani, Mohsen, and Andrew McGee. 2018. "Labor market mobility and the early-career outcomes of immigrant men." *IZA Journal of Migration and Development* 8 (1): 1–28.

Kroft, Kory, Yao Luo, Magne Mogstad, and Bradley Setzler. 2024. "Imperfect Competition and Rents in Labor and Product Markets: The Case of the Construction Industry."

Kroft, Kory, Isaac Norwich, Matthew Notowidgdo, and Stephen Tino. 2025. "The Labor Market Returns to Permanent Residency." *NBER Working Paper No. 34630.*

Lamadon, Thibaut, Magne Mogstad, and Bradley Setzler. 2022. "Imperfect Competition, Compensating Differentials, and Rent Sharing in the US Labor Market." *American Economic Review* 112 (1): 169–212.

Lehrer, Steven F., and Luke Rawling. 2025. "The Implications of Sorting for Immigrant Wage Assimilation and Changing Cohort Quality in Canada." Forthcoming in *Canadian Journal of Economics*, *NBER Working Paper No. 34462.*

Li, Jiang, Benoit Dostie, and GaÃ«lle Simard-Duplain. 2023. "Firm Pay Policies and the Gender Earnings Gap: The Mediating Role of Marital and Family Status." *ILR Review* 76 (1): 160–188.

Manning, Alan. 2003. *Monopsony in Motion: Imperfect Competition in Labor Markets.* Princeton University Press.

———. 2021. "Monopsony in Labor Markets: A Review." *ILR Review* 74 (1): 3–26.

Munshi, Kaivan. 2003. "Networks in the Modern Economy: Mexican Migrants in the U.S. Labor Market." *The Quarterly Journal of Economics* 118 (2): 549–599.

Naidu, Suresh, Yaw Nyarko, and Shing-Yi Wang. 2016. "Monopsony Power in Migrant Labor Markets: Evidence from the United Arab Emirates." *Journal of Political Economy* 124 (6): 1735–1792.

Oaxaca, Ronald. 1973. "Male-Female Wage Differentials in Urban Labor Markets." *International Economic Review* 14 (3): 693–709.

Oreopoulos, Philip. 2011. "Why Do Skilled Immigrants Struggle in the Labor Market? A Field Experiment with Thirteen Thousand Resumes." *American Economic Journal: Economic Policy* 3 (4): 148–171.

Pendakur, Krishna, and Simon Woodcock. 2010. "Glass Ceilings or Glass Doors? Wage Disparity Within and Between Firms." *Journal of Business and Economic Statistics* 28 (1): 181–189.

Robinson, Joan. 1933. *The Economics of Imperfect Competition.* London: Macmillan / Co.

Roy, Andrew Donald. 1951. "Some Thoughts on the Distribution of Earnings." *Oxford Economic Papers* 3 (2): 135–146.

Sharma, Garima. 2024. "Monopsony and Gender."

Shorrocks, Anthony F. 2013. "Decomposition Procedures for Distributional Analysis: A Unified Framework Based on the Shapley Value." *Journal of Economic Inequality* 11 (1): 99–126.

Skuterud, Mikal, and Mingcui Su. 2012. "Immigrants and the Dynamics of High-Wage Jobs." *ILR Review* 65 (2): 377–397.

Sokolova, Anna, and Todd Sorensen. 2021. "Monopsony in Labor Markets: A Meta-Analysis." *ILR Review* 74 (1): 27–55.

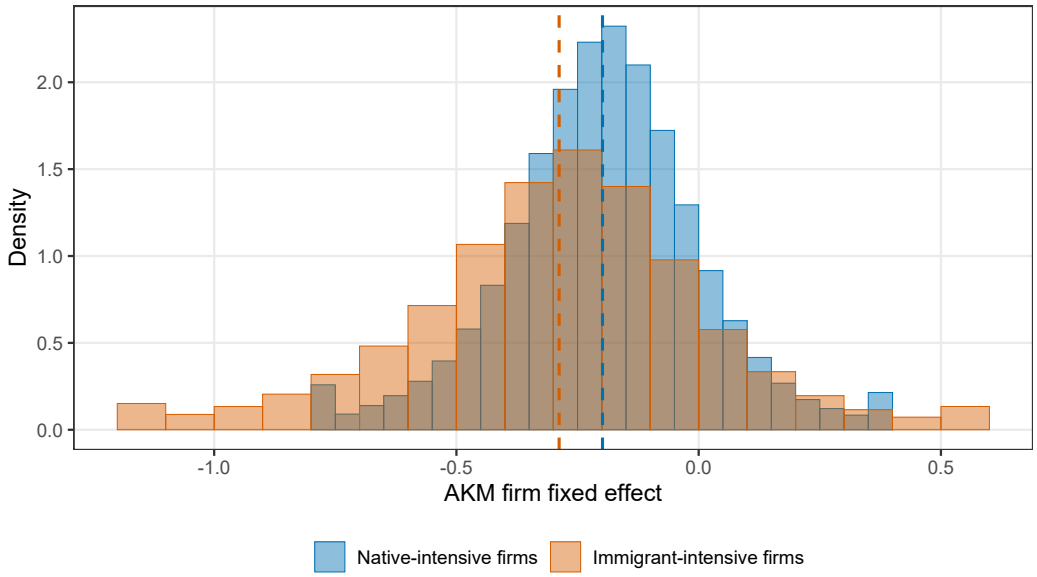
Statistics Canada. 2022. "The Daily: Immigrants make up the largest share of the population in over 150 years and continue to shape who we are as Canadians," https://www150.statcan.gc.ca/n1/daily-quotidien/221026/dq221026a-eng.htm.

Wang, Xuening. 2021. "US Permanent Residency, Job Mobility, and Earnings." *Journal of Labor Economics* 39 (3): 639–671.

Webber, DA. 2016. "Firm-Level Monopsony and the Gender Pay Gap." *Industrial Relations* 55 (2): 323–345.

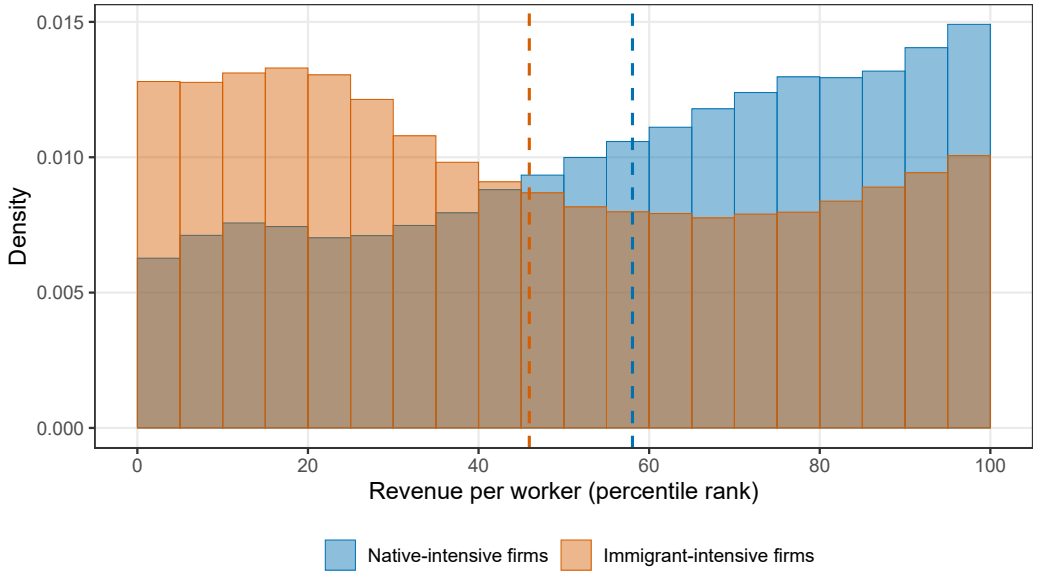
Main Figures

Figure 1: Distribution of AKM Firm Effects, by Immigrant Intensity



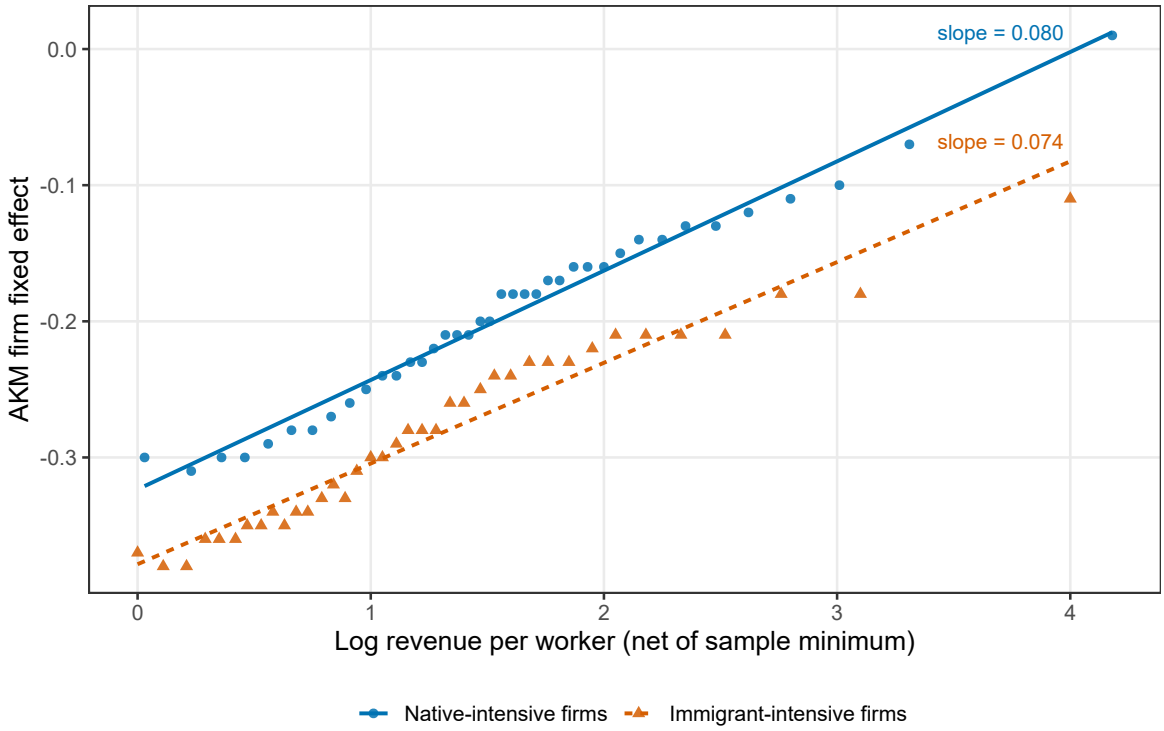
Notes: The figure plots the distribution of estimated AKM firm effects across firms, separately for immigrant-intensive and native-intensive firms. The sample is firms in the connected set that hire at least one immigrant; a firm is immigrant-intensive if its immigrant employment share is above the time-averaged median across all employers, so the two groups contain equal numbers of firms by construction. Densities are plotted rather than counts because bin widths differ across the two groups. Dashed vertical lines mark group means, which are -0.199 and -0.288. Counts underlying the histogram are rounded to comply with Statistics Canada's vetting rules. *Source:* Author's calculations using the CEEDD.

Figure 2: Distribution of Employer Revenue per Worker, by Immigrant Intensity



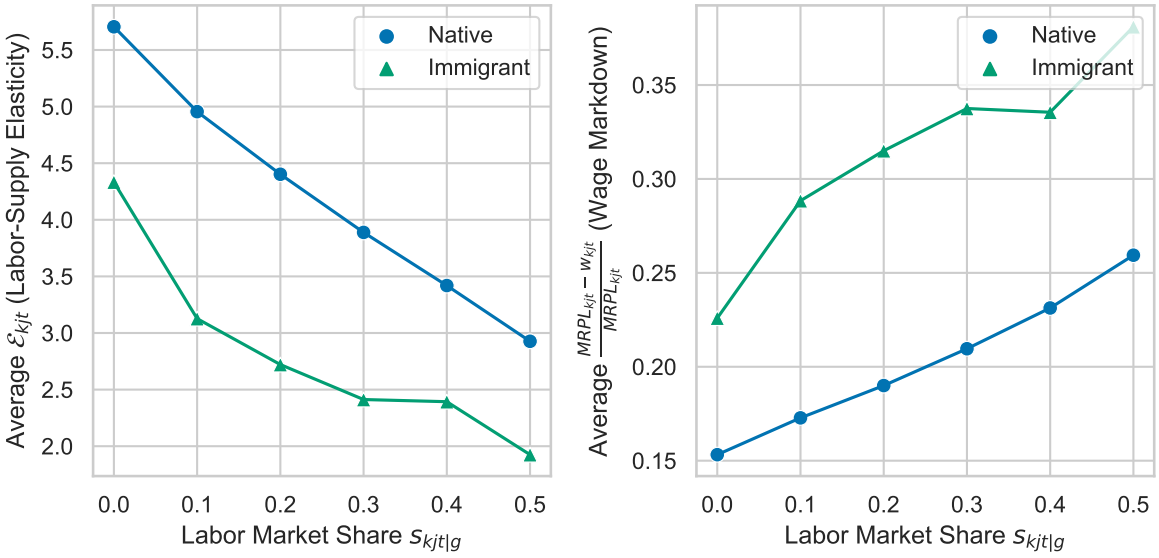
Notes: Percentile ranks of time-averaged log revenue per worker are constructed over all firms in the AKM sample. The figure then restricts to firms that hire at least one immigrant and plots the distribution of those ranks across firms, separately for immigrant-intensive and native-intensive firms, defined as in Figure 1. Both the unit of observation and the sample split are identical to Figure 1; only the variable on the horizontal axis differs. Dashed vertical lines mark group means, which are 58.0 and 45.9. Counts underlying the histogram are rounded to comply with Statistics Canada's vetting rules. *Source:* Author's calculations using the CEEDD.

Figure 3: Firm Pay Premia and Revenue per Worker, by Immigrant Intensity



Notes: This figure plots the binned relationship between the estimated AKM firm fixed effect and time-averaged log revenue per worker, with the time-averaged log revenue per worker of the bottom fifth percentile set to zero (Beauregard et al. 2025), separately for firms with an above-median immigrant employment share ("immigrant-intensive") and firms with a below-median immigrant employment share ("native-intensive"), among firms that hire at least one immigrant. Each point is a bin of the log revenue-per-worker distribution; lines are ordinary least squares fits through the binned points. The estimated slopes are 0.080 for native-intensive firms and 0.074 for immigrant-intensive firms. Conditional on log revenue per worker, immigrant-intensive firms have firm effects that are approximately 6.8 log points lower on average, compared with an unconditional difference of 8.9 log points. Appendix Figure A1 reports the analogous figure using net value added per worker. *Source:* Author's calculations using the CEEDD; revenue data come from the NALMF.

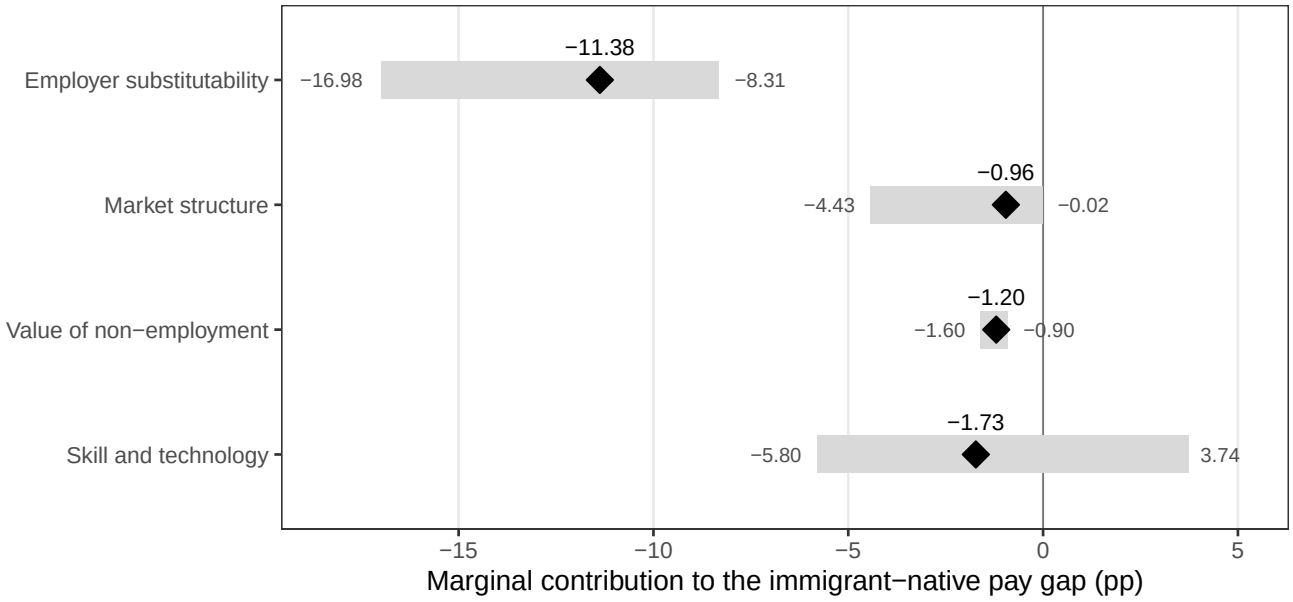
Figure 4: Labor Supply Elasticities and Markdowns by Local Labor Market Share



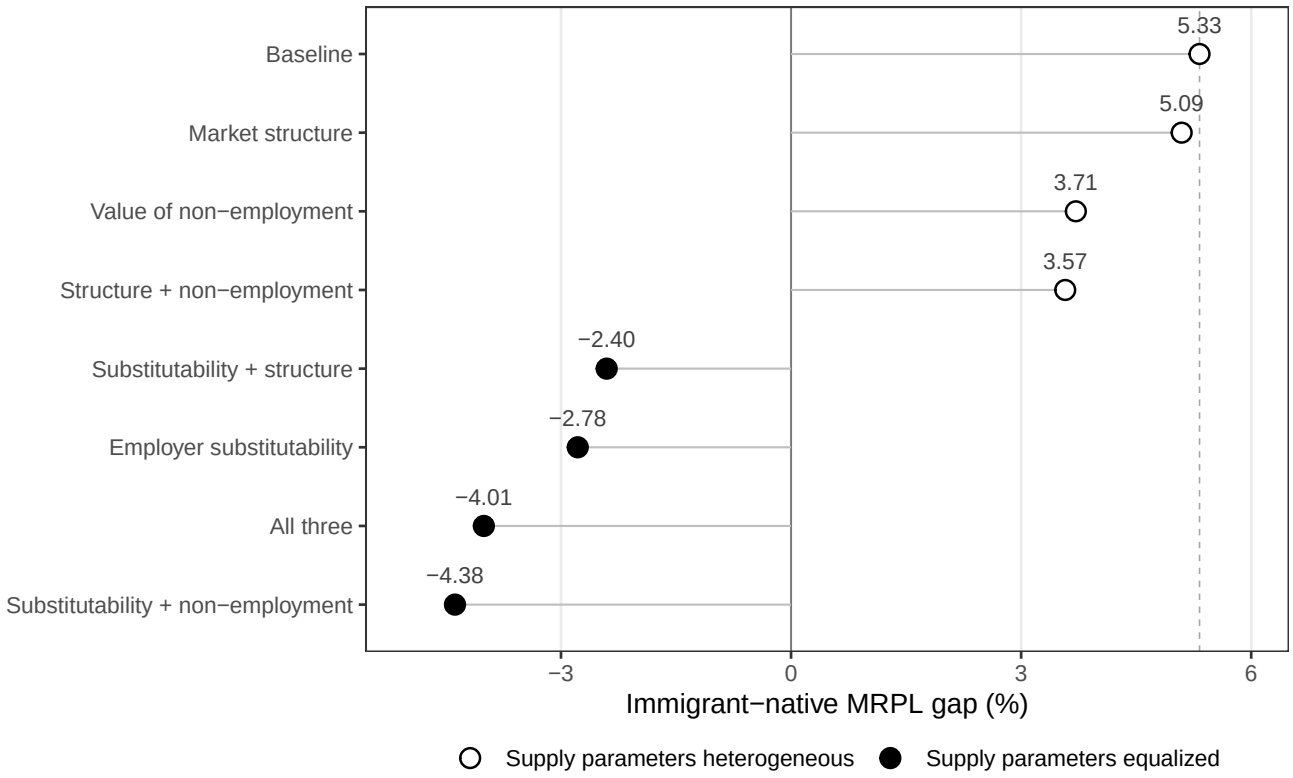
Notes: This figure illustrates the relationship between the firm-specific labor supply elasticity, wage markdown, and employer local labor market share, separately for immigrants and natives. Local labor market share, $s_{kjt|g}$, is defined as the share of type-k workers in market g employed by firm j at time t: $s_{kjt|g} \equiv l_{kjt} / \left(\sum_{j \in \mathcal{J}_g} l_{kjt} \right)$, where $\mathcal{J}_g$ is the set of firms in market g. Wage markdowns are defined as the absolute percentage below the MRPL, i.e., $(\text{MRPL}_{kjt} - w_{kjt})/\text{MRPL}_{kjt}$. *Source:* Author's calculations using the CEEDD.

Figure 5: Decomposing the Pay Gap

(a) Contribution to the pay gap



(b) MRPL gap, technology held fixed



Notes: Panel (a) plots each block's Shapley value (diamond) against the range of its eight marginal contributions (shaded bar). Panel (b) plots the immigrant-native gap in marginal revenue products across the eight counterfactuals formed from $\mathcal{S}$, $\mathcal{M}$ and $\mathcal{V}$. The technology parameters $\tilde{\theta}_j$, α_j, ρ_k and γ_{kj} are held at their estimated values in all eight, so every difference across points reflects reallocation alone. The dashed line marks the baseline gap of 5.33%. All counterfactuals use 2016 data.

Main Tables

Table 1: Summary Statistics, AKM Sample

		Ever Permanent Residents			
	Natives	All	Economic	Family	Refugee
	(1)	(2)	(3)	(4)	(5)
Panel A: Demographics					
Share male	65.0	61.5	62.3	56.3	68.8
Mean age	42.1	41.7	41.7	41.1	42.6
Share in Quebec	27.1	13.6	14.6	11.3	14.5
Share in Ontario	35.9	54.5	51.7	57.3	59.1
Share in British Columbia	11.2	15.1	15.8	16.5	9.4
Panel B: Earnings					
Mean earnings	65,500	54,500	60,300	47,800	47,300
Panel C: Employer characteristics					
Mean share immigrants at firm	11.6	51.0	49.9	53.0	51.1
Median number of workers	29	75	85	63	72
Mean log revenue	16.51	17.24	17.39	17.03	17.11
Median log revenue	16.22	17.03	17.18	16.83	16.93
Panel D: Employer revenue per worker					
Mean log revenue per worker	12.84	12.79	12.83	12.75	12.73
Median log revenue per worker	12.79	12.73	12.78	12.69	12.67
Number of person-years	74,530,000	17,610,000	9,520,000	5,400,000	2,680,000
Number of persons	10,300,000	2,950,000	1,660,000	860,000	430,000
Number of Enterprise ID's	900,000	450,000	320,000	260,000	150,000
Number of firms	2,573,000	735,000	509,000	372,000	209,000

Notes: This table contains summary statistics for the AKM sample described in Section 2. Column (1) contains native-born workers, Column (2) all individuals who are or have been permanent residents, and Columns (3)–(5) split Column (2) by admission category. Firm-level statistics are calculated as averages across the employees belonging to each column group, so that each individual is assigned the characteristics of her primary employer. All monetary units are in $2012. Earnings are rounded to the nearest $100 and counts are rounded to comply with Statistics Canada's vetting rules for intermediate output; unrounded values will be provided for publication. *Source:* Author's calculations using the CEEDD. Revenue comes from the NALMF, earnings from the T4 database, immigration status and admission category from the IMDB, and other demographic variables from the T1PMF.

Table 2: Mincer-style Regressions for Log Earnings using the 2016 Canadian Census Data

Dependent Variable:	Log Earnings				
	Baseline Specification		Additional Fixed Effects		
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
Permanent resident	-0.2075***	-0.3547***	-0.3158***	-0.3234***	-0.2733***
	(0.0057)	(0.0061)	(0.0060)	(0.0061)	(0.0060)
35-44 years		0.2250***	0.2079***	0.2061***	0.1865***
		(0.0062)	(0.0061)	(0.0062)	(0.0060)
45-54 years		0.3021***	0.2837***	0.2811***	0.2629***
		(0.0064)	(0.0062)	(0.0063)	(0.0062)
55-59 years		0.2597***	0.2436***	0.2417***	0.2269***
		(0.0083)	(0.0081)	(0.0082)	(0.0080)
Female		-0.3302***	-0.2897***	-0.2905***	-0.2654***
		(0.0048)	(0.0048)	(0.0054)	(0.0053)
Fixed-effects					
Education		Yes	Yes	Yes	Yes
Language		Yes	Yes	Yes	Yes
Marital Status		Yes	Yes	Yes	Yes
Industry			Yes		
Occupation				Yes	
Ind x Occ					Yes
Observations	197,219	197,219	197,219	197,219	197,218

Heteroskedasticity-robust standard-errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: This table reports results from the estimation of Mincer-style regressions using the Canadian 2016 Census Public-Use Microdata File (PUMF). The sample is restricted to full-time workers aged 25 to 59 with positive earnings. Temporary residents, residents of the territories, immigrants admitted before 1980 or whose admission category is not reported, and individuals with missing education data are excluded. Individuals working in the education (NAICS 61), health (NAICS 62), and public administration (NAICS 91) sectors are also excluded. Column (1) estimates a simple regression of log earnings on immigrant status without controls. Columns (2)-(5) include standard demographic controls (age and sex) alongside fixed effects for education level, primary official language, and marital status. Columns (3)-(5) introduce labor market fixed effects: column (3) includes industry fixed effects, column (4) includes occupation fixed effects, and column (5) includes interacted industry and occupation fixed effects. Heteroskedasticity-robust standard errors are in parentheses.

Table 3: Exposure to Labor Market Concentration, by Immigrant Status

	Native-born (1)	Immigrant (2)	Ratio (3) = (2)/(1)
Panel A: Collision probabilities			
Overall	0.0008	0.0010	1.33
Within-market	0.1096	0.0492	0.45
Between-market	0.0070	0.0209	2.97
Panel B: Herfindahl-Hirschman indices			
Overall	0.0008	0.0010	1.34
Within-market	0.0643	0.0482	0.75
Between-market	0.0070	0.0209	2.97

Notes: A collision probability is the probability that two workers of the indicated immigrant status, drawn at random, are employed by the same firm. It incorporates a finite-sample denominator correction for group size and decomposes multiplicatively into a within-market component (the probability that two workers employed in the same market work at the same firm) and a between-market component (the probability that two workers are employed in the same market). Panel B reports Herfindahl-Hirschman indices, computed at the market level and averaged across individuals according to the market in which they are employed; the between-market index uses market employment shares of the relevant group in place of firm shares. Labor markets are two-digit NAICS by census metropolitan area or census agglomeration, as defined in Section 2. Ratios in column (3) are computed from unrounded values and may therefore differ slightly from the ratio of the rounded entries in columns (1) and (2). *Source:* Author's calculations using the CEEDD.

Table 4: Estimated Labor Supply and Production Technology Parameters

	Labor supply		Production technology	
	Elasticity $\mathcal{E}_{kjt}$ (1)	Markdown (%) (2)	Returns to scale α_{jt} (3)	Curvature ρ_k (4)
Panel A. All workers				
All workers	5.25	17.6	0.291	0.888
	(0.08)	(0.18)	(0.001)	(0.001)
	[5.09, 5.38]	[17.2, 17.9]	[0.289, 0.292]	[0.887, 0.890]
Panel B. By immigrant status				
Native-born	5.45	16.4	0.290	0.878
	(0.08)	(0.18)	(0.001)	(0.001)
	[5.29, 5.60]	[16.1, 16.8]	[0.289, 0.292]	[0.877, 0.881]
Immigrants	4.42	22.8	0.295	0.930
	(0.16)	(0.46)	(0.001)	(0.001)
	[4.06, 4.72]	[21.6, 23.4]	[0.292, 0.296]	[0.928, 0.931]
Panel C. Immigrants, by admission category				
Economic class	5.09	20.6	0.308	0.916
	(0.22)	(0.52)	(0.001)	(0.001)
	[4.59, 5.50]	[19.4, 21.4]	[0.305, 0.309]	[0.915, 0.918]
Family class	3.85	24.6	0.279	0.947
	(0.16)	(0.62)	(0.001)	(0.001)
	[3.54, 4.16]	[23.3, 25.7]	[0.276, 0.280]	[0.945, 0.949]
Refugees	3.20	27.0	0.281	0.941
	(0.37)	(0.85)	(0.001)	(0.001)
	[2.59, 3.64]	[24.9, 28.3]	[0.278, 0.283]	[0.938, 0.942]

Notes: Employment-weighted averages of model parameters following Equation (37). Column (1) reports the firm-specific labor supply elasticity $\mathcal{E}_{kjt}$ of Equation (5), and column (2) the implied wage markdown of Equation (8), expressed as the percentage of the marginal revenue product of labor withheld by the employer. Column (3) reports returns to scale and column (4) the curvature of the technology in type-k labor, where $\rho_k = 1$ would make each type's marginal product independent of its own employment. Panel A covers the full AKM sample; Panel B partitions it into the native-born and immigrants; Panel C partitions the immigrant row of Panel B into the three admission categories, which are mutually exclusive and exhaustive within the immigrant sample. Instrumental variable estimates with bootstrap standard errors in parentheses and 95 percent bootstrap confidence intervals (Hall 1992) in brackets. In column (3) the native-born and immigrant intervals appear to share an endpoint at three decimal places but do not overlap: the native-born interval ends at 0.2915 and the immigrant interval begins at 0.2920. *Source:* Author's calculations using the CEEDD.

Table 5: Generalized Concentration Indices, by Immigrant Status

	GCI ×100 (1)	Within-market ×100 (2)	Between-market ×100 (3)
Panel A. All workers			
All workers	5.71	83.46	6.82
	(0.01)	(0.16)	(< 0.01)
	[5.70, 5.73]	[83.14, 83.76]	[6.82, 6.82]
Panel B. By immigrant status			
Native-born	4.90	84.98	5.73
	(0.01)	(0.18)	(< 0.01)
	[4.89, 4.92]	[84.60, 85.30]	[5.73, 5.73]
Immigrants	9.14	77.04	11.46
	(0.03)	(0.28)	(< 0.01)
	[9.07, 9.20]	[76.47, 77.56]	[11.46, 11.46]
Panel C. Immigrants, by admission category			
Economic class	8.69	77.35	10.89
	(0.04)	(0.36)	(< 0.01)
	[8.62, 8.77]	[76.66, 78.01]	[10.89, 10.89]
Family class	9.58	77.03	11.97
	(0.05)	(0.42)	(< 0.01)
	[9.50, 9.67]	[76.22, 77.84]	[11.97, 11.97]
Refugees	9.81	75.97	12.46
	(0.06)	(0.49)	(< 0.01)
	[9.71, 9.93]	[75.06, 76.96]	[12.46, 12.46]

Notes: Employment-weighted averages of the generalized concentration index of CKMM, defined in Section 5 and scaled by 100. Higher values denote a more concentrated environment. Column (1) reports the overall index and columns (2) and (3) its within- and between-market components. The between-market component is a function of employment shares across local labor markets only and involves no estimated parameter, so it carries no sampling uncertainty from the bootstrap; the within-market component weights each market's internal concentration by $1/\sigma_{kg}$ and inherits the sampling variation in the estimated substitution parameters. Panel A covers the full estimation sample; Panel B partitions it into the native-born and immigrants; Panel C partitions the immigrant row of Panel B into the three admission categories. Bootstrap standard errors in parentheses; 95 percent bootstrap confidence intervals (Hall 1992) in brackets. Columns (2) and (3) do not multiply to column (1), because each is an employment-weighted average of a group-year index; the decomposition is exact within each worker type and year.
Source: Author's calculations using the CEEDD.

Online Appendix for "Labor Market Power, Firm Productivity, and the Immigrant-Native Pay Gap"

A Appendix Tables

Table A1: Summary Statistics by Immigration Status in the 2016 Canadian Census, Full-time Aged 25-59 Workers Only

	Native	Ever PR
Demographics		
Male	0.644	0.604
Age 25–34 years	0.301	0.258
Age 35–44 years	0.264	0.325
Age 45–54 years	0.291	0.311
Age 55–59 years	0.144	0.106
Bachelor's degree or higher	0.201	0.438
Household size	2.889	3.626
Married/common law	0.678	0.779
Geography		
Ontario	0.338	0.511
Quebec	0.272	0.149
British Columbia	0.115	0.163
Employment Income		
Log Earnings (2012 dollars)	10.681	10.473
Industry Breakdown		
Professional, scientific & technical services	0.098	0.129
Wholesale trade	0.062	0.058
Manufacturing	0.149	0.181
Accommodation and food services	0.051	0.093
Transportation and warehousing	0.070	0.074
Retail trade	0.130	0.113
Construction	0.124	0.066
Finance, insurance and management	0.072	0.091
Admin, support, waste mgmt services	0.048	0.061
Other	0.195	0.135
N	151,193	46,026

Notes: This table presents summary statistics for full-time workers aged 25 to 59 with positive earnings by immigration status using the 2016 Canadian Census. Temporary residents, residents of the territories, immigrants admitted before 1980 or whose admission category is not reported, and individuals with missing education data are excluded. Individuals missing industry codes or working in the education (NAICS 61), health (NAICS 62), and public administration (NAICS 91) sectors are also excluded. All values except earnings and sample size are proportions. "Earnings" refers to employment income in 2012 dollars. *Source:* Authors' calculations using the 2016 Canadian Census (Public Use Microdata File).

Table A2: Summary Statistics by Immigrant Admission Class, Full-time Aged 25-59 Workers Only

	Native	Economic	Family	Refugee
Demographics				
Male	0.644	0.618	0.555	0.666
Age 25–34 years	0.301	0.268	0.245	0.246
Age 35–44 years	0.264	0.332	0.336	0.272
Age 45–54 years	0.291	0.299	0.321	0.335
Age 55–59 years	0.144	0.102	0.098	0.147
Bachelor's degree or higher	0.201	0.600	0.252	0.200
Household size	2.889	3.465	3.868	3.712
Married/common law	0.678	0.779	0.803	0.720
Geography				
Ontario	0.338	0.476	0.546	0.572
Quebec	0.272	0.158	0.125	0.168
British Columbia	0.115	0.167	0.182	0.100
Employment Income				
Log Earnings (2012 dollars)	10.681	10.585	10.331	10.348
Industry Breakdown				
Professional, scientific & technical services	0.098	0.173	0.075	0.076
Wholesale trade	0.062	0.059	0.059	0.051
Manufacturing	0.149	0.160	0.199	0.223
Accommodation and food services	0.051	0.080	0.112	0.104
Transportation and warehousing	0.070	0.058	0.099	0.081
Retail trade	0.130	0.107	0.122	0.113
Construction	0.124	0.057	0.073	0.089
Finance, insurance and management	0.072	0.110	0.071	0.059
Admin, support, waste mgmt services	0.048	0.052	0.068	0.080
Other	0.195	0.145	0.122	0.124
N	151,193	25,459	14,690	5,877

Notes: This table presents summary statistics for full-time workers aged 25 to 59 with positive earnings by immigrant admission class using the 2016 Canadian Census. Temporary residents, residents of the territories, immigrants admitted before 1980 or whose admission category is not reported, and individuals with missing education data are excluded. Individuals missing industry codes or working in the education (NAICS 61), health (NAICS 62), and public administration (NAICS 91) sectors are also excluded. All values except earnings and sample size are proportions. "Earnings" refers to employment income in 2012 dollars. *Source:* Authors' calculations using the 2016 Canadian Census (Public Use Microdata File).

Table A3: Mincer-style Regressions for Log Earnings by Admission Class

Dependent Variable:	Log Earnings				
	Baseline Specification		Additional Fixed Effects		
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
Economic Class	-0.0964***	-0.3389***	-0.3045***	-0.3181***	-0.2718***
	(0.0071)	(0.0077)	(0.0075)	(0.0076)	(0.0075)
Family Class	-0.3501***	-0.3841***	-0.3383***	-0.3415***	-0.2836***
	(0.0093)	(0.0094)	(0.0093)	(0.0094)	(0.0093)
Refugee	-0.3328***	-0.3452***	-0.3057***	-0.3008***	-0.2543***
	(0.0140)	(0.0139)	(0.0137)	(0.0137)	(0.0136)
35-44 years		0.2249***	0.2078***	0.2061***	0.1865***
		(0.0062)	(0.0061)	(0.0062)	(0.0060)
45-54 years		0.3015***	0.2833***	0.2807***	0.2627***
		(0.0064)	(0.0062)	(0.0063)	(0.0062)
55-59 years		0.2585***	0.2426***	0.2409***	0.2264***
		(0.0083)	(0.0081)	(0.0082)	(0.0080)
Female		-0.3294***	-0.2890***	-0.2900***	-0.2650***
		(0.0048)	(0.0048)	(0.0054)	(0.0053)
Fixed-effects					
Education		Yes	Yes	Yes	Yes
Language		Yes	Yes	Yes	Yes
Marital Status		Yes	Yes	Yes	Yes
Industry			Yes		
Occupation				Yes	
Ind x Occ					Yes
Observations	197,219	197,219	197,219	197,219	197,218

Heteroskedasticity-robust standard-errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: This table reports results from the estimation of Mincer-style regressions using the Canadian 2016 Census Public-Use Microdata File (PUMF). The sample is restricted to full-time workers aged 25 to 59 with positive earnings. Temporary residents, residents of the territories, immigrants admitted before 1980 or whose admission category is not reported, and individuals with missing education data are excluded. Individuals working in the education (NAICS 61), health (NAICS 62), and public administration (NAICS 91) sectors are also excluded. The omitted category is native-born workers, so each coefficient is the pay gap of that admission class relative to natives. Column (1) estimates a simple regression of log earnings on admission class without controls. Columns (2)-(5) include standard demographic controls (age and sex) alongside fixed effects for education level, primary official language, and marital status. Columns (3)-(5) introduce labor market fixed effects: column (3) includes industry fixed effects, column (4) includes occupation fixed effects, and column (5) includes interacted industry and occupation fixed effects. Heteroskedasticity-robust standard errors are in parentheses.

Table A4: Notation for the labor supply system

Object	Interpretation
β_k	Wage elasticity of employer choice at the employment margin
$\sigma_{kg} \geq 1$	Substitutability of employers within market g
$\tilde{\sigma}_{kg} = 1 - 1/\sigma_{kg}$	Coefficient on the inside share in (4)
ξ_{kjt}	Intercept: non-wage payoff of firm j to type-k workers
$s_{kjt\|g}$	Firm j's share of type-k employment within market g
$\mathcal{E}_{kjt}$	Firm-specific labor supply elasticity, Eq. (5)

Table A5: Notation for the production technology

Object	Interpretation
$C_{jt} \subset \mathcal{K}$	Set of worker types employed by firm j in period t
$\rho_k \in (0, 1]$	Curvature of the technology in type-k labor; $\rho_k = 1$ is perfect substitutability
γ_{kjt}	Match-specific productivity of type k at firm j, with $\sum_{k \in C_{jt}} \gamma_{kjt} = 1$
$\tilde{\gamma}_{kjt} = \theta_{jt}\gamma_{kjt}$	Unnormalized labor productivity; follows the AR(1) of Eq. (10)
$\alpha_{jt} \in (0, 1]$	Returns to scale of firm j in period t
$\tilde{\theta}_{jt} = \theta_{jt}^{\alpha_{jt}}$	Total factor productivity of firm j in period t
$\Lambda_{jt} = \sum_{h \in C_{jt}} \gamma_{hjt} l_{hjt}^{\rho_h}$	Effective labor aggregate, so that $R_{jt} = \tilde{\theta}_{jt}\Lambda_{jt}^{\alpha_{jt}}$
δ	Persistence of the labor productivity process, common across types
MRPL_{kjt}	Marginal revenue product of type-k labor, Eq. (7)
$1/(1 + \mathcal{E}_{kjt})$	Wage markdown, Eq. (8)

Table A6: Summary Statistics, Estimation Sample

		Ever Permanent Residents			
	Natives	All	Economic	Family	Refugee
	(1)	(2)	(3)	(4)	(5)
Panel A: Demographics					
Share male	63.7	61.1	61.9	56.0	68.5
Mean age	42.3	41.8	41.8	41.3	42.7
Share in Quebec	27.4	13.6	14.7	11.3	14.5
Share in Ontario	36.5	55.2	52.3	58.2	59.6
Share in British Columbia	11.0	14.7	15.5	16.0	9.4
Panel B: Earnings					
Mean earnings	67,000	55,000	61,000	48,000	48,000
Panel C: Employer characteristics					
Mean share immigrants at firm	13.4	48.4	47.1	50.5	48.6
Median number of workers	47	96	112	80	86
Mean log revenue	17.06	17.61	17.79	17.39	17.43
Median log revenue	16.74	17.31	17.50	17.10	17.15
Panel D: Employer revenue per worker					
Mean log revenue per worker	12.90	12.84	12.88	12.79	12.77
Median log revenue per worker	12.86	12.78	12.83	12.73	12.72
Panel E: Retention relative to the AKM sample					
Share of person-years retained	74.4	81.5	81.3	81.6	82.1
Share of persons retained	80.2	83.5	83.7	82.8	84.2
Number of person-years	55,467,000	14,347,000	7,742,000	4,406,000	2,199,000
Number of persons	8,264,000	2,463,000	1,389,000	712,000	362,000
Number of Enterprise ID's	270,000	172,000	138,000	119,000	78,000
Number of firms	379,000	233,000	186,000	156,000	100,000

Notes: This table replicates Table 1 for the estimation sample of Section 4.6, which is the subset of the AKM sample on which Equation (12) can be evaluated. Column definitions and construction are identical to Table 1. Panel E expresses the counts in the final rows as shares of the corresponding counts in Table 1. Earnings are rounded to the nearest $1,000 and counts are rounded to comply with Statistics Canada's vetting rules for intermediate output; unrounded values will be provided for publication.
Source: Author's calculations using the CEEDD.

Table A7: Counterfactual blocks

Block	Name	Parameters restricted	Policy lever	Direct effect
$\mathcal{S}$	Employer substitutability	β_k, σ_{kg}	Information and mobility	Markdown
$\mathcal{M}$	Market structure	No strategic interactions in wage-setting	Antitrust; non-competes	Markdown
$\mathcal{V}$	Value of non-employment	w_{0k}	Settlement supports; social assistance	None
$\mathcal{R}$	Skill mix and technology	m_k, ρ_k, $\tilde{\theta}_j$, α_j, γ_{kj}, ξ_{kj}	Selection criteria; training; industrial policy	MRPL

Notes: The blocks are exhaustive over the estimated primitives, so the Shapley decomposition of Section 6.3 allocates the entire gap. Each block is defined so as to correspond to a distinct policy lever, listed in column four; the decomposition therefore reports how much of the gap each lever can reach. Within $\mathcal{R}$, m_k and ρ_k are properties of workers, $\tilde{\theta}_j$ and α_j of firms, and γ_{kj} and ξ_{kj} of worker-firm pairs; the last two are grouped here rather than with $\mathcal{S}$ because they govern what a type can produce and where, and because they are not separately identified from one another. "Direct effect" follows the definition in Section 6.1: the object a manipulation moves before any reallocation of workers across firms. Every block also has indirect effects through general-equilibrium propagation.

Table A8: First-Stage Partial F-Statistics

	β_k	σ_{kg}, by province group			
		British Columbia	Ontario	Quebec	Other
	(1)	(2)	(3)	(4)	(5)
Native-born					
Women	1,171	12,040	13,351	10,776	12,730
Men	1,972	15,087	16,365	13,571	17,089
Economic class					
Americas, Women	12	330	309	412	374
Europe, Women	54	942	950	851	819
Africa, Women	10	261	209	214	163
Asia, Women	69	1,869	1,845	1,451	1,432
Americas, Men	38	598	468	703	712
Europe, Men	131	1,420	1,547	1,526	1,659
Africa, Men	32	526	462	480	567
Asia, Men	144	2,873	2,826	2,553	2,557
Family class					
Americas, Women	22	360	372	245	336
Europe, Women	57	335	442	501	332
Africa, Women	5	94	89	57	97
Asia, Women	115	1,068	1,099	915	924
Americas, Men	88	574	616	605	767
Europe, Men	75	711	717	732	820
Africa, Men	15	142	152	141	186
Asia, Men	191	1,385	1,630	1,300	1,915
Refugees					
Americas, Women	9	111	141	116	116
Europe, Women	25	271	268	200	248
Africa, Women	4	113	60	55	91
Asia, Women	41	382	363	359	361
Americas, Men	38	268	315	189	390
Europe, Men	76	451	531	461	583
Africa, Men	10	234	181	254	181
Asia, Men	125	809	908	826	877

Notes: Partial F-statistics from the first stage of the estimation of Equation (12). Column (1) reports the statistic for β_k and columns (2) to (5) for σ_{kg}, which is assumed constant within each combination of worker type and province group (British Columbia, Ontario, Quebec, and all other provinces). Worker types are the 26 k-groups defined in Section 4. Three of the 26 types have a first-stage F-statistic below 10 for β_k—family-class and refugee women born in Africa, and refugee women born in the Americas—together accounting for less than 1 percent of the full AKM sample and less than 3 percent of immigrants in the AKM sample; the main results for the immigrant-native pay gap are similar when these types are dropped or pooled with adjacent categories. Every statistic for σ_{kg} exceeds 50.
Source: Author's calculations using the CEEDD.

Table A9: Primitives of the Firm-Specific Labor Supply Elasticity

	β_k (1)	σ_{kg} (2)	$\beta_k\sigma_{kg}$ (3)	$\mathcal{E}_{kjt}$ (4)
Panel A. All workers				
All workers	0.559	10.79	6.03	5.25
	(0.006)	(0.15)		(0.08)
	[0.549, 0.572]	[10.45, 11.03]		[5.09, 5.38]
Panel B. By immigrant status				
Native-born	0.526	11.73	6.17	5.45
	(0.006)	(0.18)		(0.08)
	[0.515, 0.538]	[11.34, 12.02]		[5.29, 5.60]
Immigrants	0.697	6.81	4.75	4.42
	(0.017)	(0.16)		(0.16)
	[0.671, 0.736]	[6.44, 7.05]		[4.06, 4.72]
Panel C. Immigrants, by admission category				
Economic class	0.725	7.57	5.49	5.09
	(0.023)	(0.23)		(0.22)
	[0.681, 0.775]	[7.02, 7.97]		[4.59, 5.50]
Family class	0.636	6.39	4.06	3.85
	(0.019)	(0.18)		(0.16)
	[0.601, 0.676]	[5.98, 6.67]		[3.54, 4.16]
Refugees	0.723	4.97	3.59	3.20
	(0.031)	(0.33)		(0.37)
	[0.673, 0.792]	[4.41, 5.28]		[2.59, 3.64]

Notes: Employment-weighted averages following Equation (37). Column (1) reports β_k, the elasticity of employer choice with respect to pay at the employment margin. Column (2) reports σ_{kg}, governing substitutability among employers within a local labor market. Column (3) reports the product of columns (1) and (2), which is the labor supply elasticity facing an employer with negligible market share; it is computed from the reported averages and is therefore not itself an employment-weighted average, so no standard error is reported. Column (4) reproduces the realized firm-specific elasticity from Table 4. The gap between columns (3) and (4) reflects the size-dependent term of Equation (5). Panel A covers the full AKM sample; Panel B partitions it into the native-born and immigrants; Panel C partitions the immigrant row of Panel B into the three admission categories. Bootstrap standard errors in parentheses and 95 percent bootstrap confidence intervals (Hall 1992) in brackets.
Source: Author's calculations using the CEEDD.

Table A10: Employer-Choice Elasticity β_k, by Worker Type

	Women (1)	Men (2)
Native-born		
All	0.708 (0.010)	0.427 (0.005)
Economic class		
Americas	1.312 (0.163)	0.756 (0.061)
Europe	0.692 (0.049)	0.452 (0.021)
Africa	0.873 (0.148)	0.572 (0.059)
Asia	1.001 (0.065)	0.626 (0.026)
Family class		
Americas	1.283 (0.117)	0.713 (0.033)
Europe	0.472 (0.041)	0.379 (0.025)
Africa	0.712 (0.191)	0.625 (0.083)
Asia	0.697 (0.035)	0.499 (0.019)
Refugees		
Americas	1.260 (0.196)	0.675 (0.052)
Europe	0.581 (0.074)	0.372 (0.023)
Africa	1.443 (0.327)	1.354 (0.178)
Asia	0.680 (0.064)	0.606 (0.026)

Notes: Worker types are the 26 k-groups defined in Section 4: native-born workers by gender, and permanent residents by admission category, macroregion of birth, and gender. "Europe" additionally includes the United States, Australia, and New Zealand. Bootstrap standard errors in parentheses. β_k is the elasticity of employer choice with respect to pay at the employment margin, estimated from Equation (4). Aggregates are in Table A9.
Source: Author's calculations using the CEEDD.

Table A11: Within-Market Substitutability σ_{kg}, by Worker Type

	Women (1)	Men (2)
Native-born		
All	7.63 (0.12)	13.94 (0.25)
Economic class		
Americas	3.60 (0.24)	5.51 (0.44)
Europe	4.13 (0.16)	6.12 (0.23)
Africa	3.33 (0.27)	6.26 (0.57)
Asia	7.17 (0.35)	10.32 (0.55)
Family class		
Americas	5.22 (0.50)	7.14 (0.49)
Europe	3.08 (0.13)	4.59 (0.22)
Africa	3.16 (0.85)	5.21 (0.97)
Asia	7.28 (0.37)	7.53 (0.33)
Refugees		
Americas	3.54 (6.06)	4.61 (0.33)
Europe	3.21 (0.32)	4.33 (0.20)
Africa	2.87 (0.52)	4.32 (0.80)
Asia	4.77 (0.28)	6.46 (0.31)

Notes: Worker types are the 26 k-groups defined in Section 4: native-born workers by gender, and permanent residents by admission category, macroregion of birth, and gender. "Europe" additionally includes the United States, Australia, and New Zealand. Bootstrap standard errors in parentheses. σ_{kg} governs substitutability among employers within a local labor market; larger values denote employers that are closer substitutes for the group. Entries are employment-weighted averages across the local labor markets in which the type is observed. Aggregates are in Table A9.
Source: Author's calculations using the CEEDD.

Table A12: Firm-Specific Labor Supply Elasticity $\mathcal{E}_{kjt}$, by Worker Type

	Women (1)	Men (2)
Native-born		
All	5.09 (0.09)	5.63 (0.10)
Economic class		
Americas	4.42 (0.60)	3.91 (0.47)
Europe	2.71 (0.20)	2.64 (0.16)
Africa	2.71 (0.50)	3.33 (0.46)
Asia	6.82 (0.51)	6.16 (0.40)
Family class		
Americas	6.29 (0.84)	4.85 (0.42)
Europe	1.36 (0.12)	1.64 (0.13)
Africa	2.06 (0.72)	2.95 (0.66)
Asia	4.79 (0.33)	3.59 (0.20)
Refugees		
Americas	4.01 (6.43)	2.90 (0.33)
Europe	1.71 (0.27)	1.50 (0.12)
Africa	3.65 (1.07)	5.16 (1.37)
Asia	3.01 (0.32)	3.72 (0.25)

Notes: Worker types are the 26 k-groups defined in Section 4: native-born workers by gender, and permanent residents by admission category, macroregion of birth, and gender. "Europe" additionally includes the United States, Australia, and New Zealand. Bootstrap standard errors in parentheses. Entries are employment-weighted averages of the firm-specific labor supply elasticity of Equation (5). Aggregates are in Table 4.
Source: Author's calculations using the CEEDD.

Table A13: Wage Markdowns, by Worker Type

	Women (1)	Men (2)
Native-born		
All	17.2 (0.2)	16.0 (0.2)
Economic class		
Americas	19.6 (2.1)	22.7 (1.8)
Europe	28.2 (1.4)	29.6 (1.1)
Africa	27.9 (3.9)	25.6 (2.1)
Asia	14.6 (0.8)	17.4 (0.7)
Family class		
Americas	15.9 (1.4)	19.0 (1.1)
Europe	43.4 (2.2)	39.8 (1.8)
Africa	34.3 (6.4)	28.2 (3.5)
Asia	18.8 (0.9)	23.0 (0.9)
Refugees		
Americas	21.4 (3.1)	27.1 (2.0)
Europe	38.0 (3.2)	40.8 (1.9)
Africa	22.9 (4.9)	18.8 (2.6)
Asia	26.1 (2.0)	22.5 (1.0)

Notes: Worker types are the 26 k-groups defined in Section 4: native-born workers by gender, and permanent residents by admission category, macroregion of birth, and gender. "Europe" additionally includes the United States, Australia, and New Zealand. Bootstrap standard errors in parentheses. Entries are the percentage of the marginal revenue product of labor withheld by the employer, from Equation (8). Aggregates are in Table 4.
Source: Author's calculations using the CEEDD.

Table A14: Labor Supply Estimates: Instrumental Variables versus Ordinary Least Squares

	Elasticity $\mathcal{E}_{kjt}$		Markdown (%)	
	IV (1)	OLS (2)	IV (3)	OLS (4)
Panel A. All workers				
All workers	5.25	0.289	17.6	78.5
	(0.08)	(0.001)	(0.18)	(0.07)
Panel B. By immigrant status				
Native-born	5.45	0.341	16.4	74.6
	(0.08)	(0.001)	(0.18)	(0.07)
Immigrants	4.42	0.069	22.8	94.7
	(0.16)	(0.002)	(0.46)	(0.17)
Panel C. Immigrants, by admission category				
Economic class	5.09	0.159	20.6	86.5
	(0.22)	(0.002)	(0.52)	(0.16)
Family class	3.85	−0.057	24.6	—
	(0.16)	(0.002)	(0.62)	
Refugees	3.20	0.005	27.0	100.1
	(0.37)	(0.002)	(0.85)	(0.22)

Notes: Employment-weighted averages following Equation (37). Columns (1) and (3) reproduce the instrumental variable estimates of Table 4; columns (2) and (4) report estimates of the same specification, Equation (4), by ordinary least squares. The markdown is the percentage of the marginal revenue product of labor withheld by the employer. Panel A covers the full AKM sample; Panel B partitions it into the native-born and immigrants; Panel C partitions the immigrant row of Panel B into the three admission categories, which are mutually exclusive and exhaustive within the immigrant sample. Bootstrap standard errors in parentheses. The least-squares elasticity for family-class immigrants is negative and therefore lies outside the parameter space implied by Assumption 4.1, so the corresponding markdown is undefined; the least-squares estimate for refugees implies a markdown of essentially 100 percent.
Source: Author's calculations using the CEEDD.

Table A15: Production Technology Parameters, Including Jointly Identified Productivity Terms

	α_{jt} (1)	ρ_k (2)	$\log \tilde{\theta}_{jt}$ (3)	γ_{kjt} (4)
Panel A. All workers				
All workers	0.2912	0.8882	15.947	0.3759
	(0.0008)	(0.0009)	(0.002)	(0.0004)
	[0.2893, 0.2923]	[0.8866, 0.8900]	[15.944, 15.950]	[0.3753, 0.3770]
Panel B. By immigrant status				
Native-born	0.2903	0.8785	15.815	0.4240
	(0.0008)	(0.0010)	(0.002)	(0.0006)
	[0.2886, 0.2915]	[0.8767, 0.8805]	[15.812, 15.818]	[0.4232, 0.4254]
Immigrants	0.2948	0.9295	16.505	0.1724
	(0.0011)	(0.0007)	(0.002)	(0.0002)
	[0.2920, 0.2961]	[0.9282, 0.9305]	[16.502, 16.508]	[0.1719, 0.1728]
Panel C. Immigrants, by admission category				
Economic class	0.3076	0.9164	16.591	0.1732
	(0.0011)	(0.0009)	(0.002)	(0.0004)
	[0.3048, 0.3090]	[0.9146, 0.9181]	[16.587, 16.594]	[0.1725, 0.1740]
Family class	0.2789	0.9472	16.375	0.1757
	(0.0011)	(0.0008)	(0.002)	(0.0006)
	[0.2762, 0.2803]	[0.9454, 0.9486]	[16.373, 16.378]	[0.1743, 0.1767]
Refugees	0.2814	0.9408	16.460	0.1628
	(0.0012)	(0.0010)	(0.002)	(0.0008)
	[0.2784, 0.2829]	[0.9379, 0.9422]	[16.457, 16.463]	[0.1609, 0.1642]

Notes: Columns (1) and (2) reproduce Table 4 at a higher precision. Column (3) reports log total factor productivity and column (4) match-specific productivity. These two parameters enter Equation (6) only through the product $\tilde{\theta}_{jt}\gamma_{kjt}$ and are separated by the normalization $\sum_{k \in C_{jt}} \gamma_{kjt} = 1$, so their levels are not comparable across groups that differ in the number of worker types their employers draw on; see Section 5.2 and Appendix I.3. Panel A covers the full AKM sample; Panel B partitions it into the native-born and immigrants; Panel C partitions the immigrant row of Panel B into the three admission categories. Bootstrap standard errors in parentheses and 95 percent bootstrap confidence intervals (Hall 1992) in brackets.

Source: Author's calculations using the CEEDD.

Table A16: Returns to Scale α_{jt}, by Worker Type

	Mean		Median	
	Women (1)	Men (2)	Women (3)	Men (4)
Native-born				
All	0.291 (0.001)	0.290 (0.001)	0.223 (0.001)	0.241 (0.001)
Economic class				
Americas	0.303 (0.001)	0.307 (0.001)	0.240 (0.001)	0.251 (0.001)
Europe	0.320 (0.001)	0.336 (0.002)	0.261 (0.001)	0.277 (0.001)
Africa	0.315 (0.002)	0.322 (0.002)	0.247 (0.002)	0.252 (0.002)
Asia	0.292 (0.001)	0.301 (0.001)	0.215 (0.001)	0.228 (0.001)
Family class				
Americas	0.279 (0.001)	0.267 (0.001)	0.216 (0.001)	0.226 (0.001)
Europe	0.311 (0.003)	0.323 (0.003)	0.256 (0.002)	0.271 (0.002)
Africa	0.292 (0.005)	0.293 (0.002)	0.230 (0.004)	0.234 (0.002)
Asia	0.269 (0.001)	0.265 (0.001)	0.206 (0.001)	0.215 (0.001)
Refugees				
Americas	0.286 (0.001)	0.286 (0.001)	0.232 (0.001)	0.242 (0.001)
Europe	0.301 (0.003)	0.312 (0.002)	0.250 (0.002)	0.272 (0.002)
Africa	0.298 (0.002)	0.300 (0.002)	0.233 (0.002)	0.236 (0.001)
Asia	0.268 (0.001)	0.260 (0.001)	0.214 (0.001)	0.219 (0.001)

Notes: Worker types are the 26 k-groups defined in Section 4: native-born workers by gender, and permanent residents by admission category, macroregion of birth, and gender. "Europe" additionally includes the United States, Australia, and New Zealand. Bootstrap standard errors in parentheses. α_{jt} is a firm-level parameter: each worker is assigned her employer's value, and the mean and median are taken across workers of that type. Values below one denote decreasing returns to scale. Aggregates are in Table 4.
Source: Author's calculations using the CEEDD.

Table A17: Curvature of the Technology ρ_k, by Worker Type

	Baseline		Alternative	
	Women (1)	Men (2)	Women (3)	Men (4)
Native-born				
All	0.946 (0.001)	0.842 (0.001)	0.928 (0.001)	0.797 (0.001)
Economic class				
Americas	0.925 (0.004)	0.923 (0.003)	0.910 (0.004)	0.902 (0.003)
Europe	0.902 (0.002)	0.908 (0.001)	0.875 (0.002)	0.878 (0.002)
Africa	0.929 (0.005)	0.947 (0.003)	0.917 (0.006)	0.941 (0.004)
Asia	0.918 (0.002)	0.915 (0.001)	0.900 (0.002)	0.891 (0.002)
Family class				
Americas	0.911 (0.003)	0.948 (0.003)	0.897 (0.004)	0.940 (0.004)
Europe	0.947 (0.003)	0.938 (0.002)	0.937 (0.003)	0.923 (0.003)
Africa	0.941 (0.008)	0.957 (0.004)	0.939 (0.009)	0.955 (0.005)
Asia	0.961 (0.001)	0.948 (0.001)	0.960 (0.002)	0.938 (0.001)
Refugees				
Americas	0.908 (0.004)	0.926 (0.002)	0.892 (0.004)	0.913 (0.002)
Europe	0.959 (0.004)	0.929 (0.004)	0.950 (0.005)	0.910 (0.005)
Africa	0.938 (0.005)	0.941 (0.004)	0.933 (0.005)	0.932 (0.004)
Asia	0.979 (0.002)	0.935 (0.001)	0.981 (0.003)	0.922 (0.001)

Notes: Worker types are the 26 k-groups defined in Section 4: native-born workers by gender, and permanent residents by admission category, macroregion of birth, and gender. "Europe"additionally includes the United States, Australia, and New Zealand. Bootstrap standard errors in parentheses. ρ_k governs the rate at which the marginal product of type-k labor declines in type-k employment; $\rho_k = 1$ would make types perfect substitutes. The upper bound of the 95 percent bootstrap confidence interval lies below unity for every type, the largest being 0.982. Aggregates are in Table 4.

Source: Author's calculations using the CEEDD.

Table A18: Log Total Factor Productivity $\log \tilde{\theta}_{jt}$, by Worker Type

	Mean		Median	
	Women (1)	Men (2)	Women (3)	Men (4)
Native-born				
All	15.860 (0.002)	15.791 (0.002)	13.607 (<0.001)	13.360 (<0.001)
Economic class				
Americas	17.011 (0.002)	16.769 (0.002)	15.580 (0.001)	15.395 (0.001)
Europe	16.361 (0.002)	16.131 (0.002)	14.828 (0.001)	14.654 (0.001)
Africa	16.795 (0.002)	16.731 (0.002)	15.424 (0.001)	15.250 (0.001)
Asia	16.696 (0.002)	16.671 (0.002)	14.548 (0.001)	14.401 (0.001)
Family class				
Americas	16.927 (0.002)	16.501 (0.002)	15.409 (0.001)	15.195 (0.001)
Europe	16.079 (0.001)	15.831 (0.001)	14.876 (0.001)	14.840 (0.001)
Africa	16.810 (0.002)	16.631 (0.002)	15.736 (0.002)	15.670 (0.001)
Asia	16.447 (0.002)	16.321 (0.002)	14.582 (0.001)	14.596 (0.001)
Refugees				
Americas	16.457 (0.002)	16.189 (0.001)	15.613 (0.001)	15.379 (0.001)
Europe	16.484 (0.002)	16.276 (0.002)	15.355 (0.001)	15.130 (0.001)
Africa	16.908 (0.003)	16.831 (0.002)	15.888 (0.001)	15.673 (0.001)
Asia	16.660 (0.002)	16.373 (0.001)	15.173 (0.001)	14.812 (0.001)

Notes: Worker types are the 26 k-groups defined in Section 4: native-born workers by gender, and permanent residents by admission category, macroregion of birth, and gender. "Europe"additionally includes the United States, Australia, and New Zealand. Bootstrap standard errors in parentheses. $\tilde{\theta}_{jt}$ is a firm-level parameter: each worker is assigned her employer's value, and the mean and median are taken across workers of that type. $\tilde{\theta}_{jt}$ and γ_{kjt} are identified only jointly, up to the normalization $\sum_{k \in C_{jt}} \gamma_{kjt} = 1$, so levels are not comparable across types whose employers draw on different numbers of worker types; see Appendix I.3. Aggregates are in Table A15.

Source: Author's calculations using the CEEDD.

Table A19: Match-Specific Productivity γ_{kjt}, by Worker Type

	Women (1)	Men (2)
Native-born		
All	0.284 (0.001)	0.499 (0.001)
Economic class		
Americas	0.108 (0.001)	0.146 (0.002)
Europe	0.179 (0.002)	0.245 (0.002)
Africa	0.144 (0.004)	0.159 (0.002)
Asia	0.141 (0.001)	0.176 (0.001)
Family class		
Americas	0.104 (0.001)	0.143 (0.001)
Europe	0.222 (0.004)	0.279 (0.004)
Africa	0.128 (0.010)	0.151 (0.004)
Asia	0.141 (0.001)	0.196 (0.001)
Refugees		
Americas	0.118 (0.003)	0.167 (0.003)
Europe	0.163 (0.005)	0.230 (0.004)
Africa	0.089 (0.004)	0.109 (0.002)
Asia	0.126 (0.002)	0.178 (0.001)

Notes: Worker types are the 26 k-groups defined in Section 4: native-born workers by gender, and permanent residents by admission category, macroregion of birth, and gender. "Europe" additionally includes the United States, Australia, and New Zealand. Bootstrap standard errors in parentheses. Entries are employment-weighted averages of γ_{kjt} across the employers at which the type is observed. γ_{kjt} and $\tilde{\theta}_{jt}$ are identified only jointly, up to the normalization $\sum_{k \in C_{jt}} \gamma_{kjt} = 1$, so levels are not comparable across types whose employers draw on different numbers of worker types; see Appendix I.3. Aggregates are in Table A15.
Source: Author's calculations using the CEEDD.

Table A20: Generalized Concentration Indices, by Worker Type

	GCI		Within-market		Between-market	
	Women (1)	Men (2)	Women (3)	Men (4)	Women (5)	Men (6)
Native-born						
All	10.32 (0.02)	1.99 (0.01)	86.01 (0.18)	84.42 (0.24)	11.99 (<0.01)	2.35 (<0.01)
Economic class						
Americas	12.32 (0.22)	3.34 (0.09)	74.57 (1.32)	67.71 (1.74)	16.49 (<0.01)	4.92 (<0.01)
Europe	9.88 (0.10)	2.76 (0.03)	74.61 (0.77)	68.11 (0.79)	13.22 (<0.01)	4.03 (<0.01)
Africa	19.73 (0.31)	6.83 (0.11)	80.59 (1.26)	76.87 (1.24)	24.47 (<0.01)	8.87 (<0.01)
Asia	15.76 (0.10)	6.30 (0.05)	84.01 (0.55)	78.90 (0.65)	18.70 (<0.01)	7.89 (<0.01)
Family class						
Americas	16.02 (0.24)	4.33 (0.08)	80.56 (1.20)	71.78 (1.35)	19.88 (<0.01)	6.03 (<0.01)
Europe	12.09 (0.14)	2.74 (0.04)	75.18 (0.84)	66.30 (1.09)	16.07 (<0.01)	4.12 (<0.01)
Africa	29.03 (0.37)	5.73 (0.17)	86.21 (1.11)	75.97 (2.23)	33.67 (<0.01)	7.52 (<0.01)
Asia	16.01 (0.10)	5.44 (0.06)	85.22 (0.55)	75.09 (0.81)	18.78 (<0.01)	7.24 (<0.01)
Refugees						
Americas	17.50 (0.29)	3.72 (0.09)	82.30 (1.37)	68.85 (1.60)	21.24 (<0.01)	5.40 (<0.01)
Europe	10.95 (0.16)	4.46 (0.07)	75.33 (1.13)	69.28 (1.11)	14.54 (<0.01)	6.44 (<0.01)
Africa	31.63 (0.44)	7.67 (0.21)	88.47 (1.24)	77.76 (2.17)	35.73 (<0.01)	9.81 (<0.01)
Asia	18.24 (0.17)	7.29 (0.09)	83.26 (0.78)	75.23 (0.89)	21.91 (<0.01)	9.69 (<0.01)

Notes: Worker types are the 26 k-groups defined in Section 4: native-born workers by gender, and permanent residents by admission category, macroregion of birth, and gender. "Europe" additionally includes the United States, Australia, and New Zealand. Bootstrap standard errors in parentheses. All indices are scaled by 100. Higher values denote a more concentrated environment; expected worker welfare is decreasing in the overall index. The between-market component depends only on the distribution of employment across local labor markets and involves no estimated parameter, so it carries no sampling uncertainty from the bootstrap. The within- and between-market components do not multiply to the overall index, because each is an employment-weighted average of a type-year index. Aggregates are in Table 5.
Source: Author's calculations using the CEEDD.

Table A21: Correlations of estimated amenities with illness or injury

Dependent variable: $\hat{\psi}^u_{kn}$		
	(1)	(2)
Cases	-0.066	-0.072
	(0.028)	(0.098)
Immigrant		0.411
		(0.306)
Immigrant × Cases		0.007
		(0.103)
Observations	520	520
R-squared	0.01	0.022

This table presents results from the estimation of Equation (47). The dependent variable is $\hat{\psi}^u_{kn}$, which are the estimated industry fixed effects from the regression of the firm-specific supply curve intercept on firm characteristics (see Equation (46)). The RHS variable "Cases" refers to the cases of illness or injury per 100,000 people. Column (1) shows the simple linear regression of Equation (47). Column (2) shows the results from a similar regression model that includes an interaction of cases with immigrant status. *Source:* Author's calculations from the Canadian Employer-Employee Dynamics Database (CEEDD) and The U.S. Bureau of Labor Statistics Data on Injury, Illness, and Fatalities.

Table A22: Labor Supply Elasticities and Wage Markdowns by Province Group

	Elasticity $\mathcal{E}_{kjt}$		Markdown (%)	
	Native-born (1)	Immigrants (2)	Native-born (3)	Immigrants (4)
Ontario	6.03	5.16	15	19
British Columbia	5.81	5.14	15	20
Quebec	5.96	3.72	15	25
Other provinces	3.94	1.95	21	36

Notes: Employment-weighted averages following Equation (37), aggregated to the four province groups over which σ_{kg} is permitted to vary. The markdown is the percentage of the marginal revenue product of labor withheld by the employer. "Other provinces" comprises the Prairies and Atlantic Canada.
Source: Author's calculations using the CEEDD.

Table A23: Order-averaged (Shapley) contributions

Block	Name and policy lever	Shapley (pp)	Shapley (%)	Range over orders Min	Range over orders Max	Sign stable
Panel A: Wage-setting power						
$\mathcal{S}$	Employer substitutability	−11.38	74.6	−16.98	−8.31	Yes
	Information and mobility					
$\mathcal{M}$	Market structure	−0.96	6.3	−4.43	−0.02	Yes
	Antitrust; non-competes					
Subtotal		**−12.33**	**80.8**			
Panel B: Worker and firm fundamentals						
$\mathcal{V}$	Value of non-employment	−1.20	7.9	−1.60	−0.90	Yes
	Settlement supports; social assistance					
$\mathcal{R}$	Skill mix and technology	−1.73	11.3	−5.80	+3.74	**No**
	Selection criteria; industrial policy					
Subtotal		**−2.93**	**19.2**			
Total		**−15.26**	**100.0**			

Notes: Shapley values from Equation (19), computed on 2016 data. Blocks and their policy levers are defined in Table A7. Panel A collects the two blocks that act on the markdown and therefore on the share of the marginal revenue product a worker is paid; Panel B collects the two that act on what a worker produces and on the value of the outside option. "(%)" is the contribution as a share of the total change, $\mathcal{D}(\mathcal{B}) - \mathcal{D}(\emptyset) = -15.26$ pp; exact additivity guarantees that the shares sum to one, though the rounded entries shown need not sum exactly to the subtotals. "Range" is the minimum and maximum marginal contribution across the eight subsets of the remaining blocks; "sign stable" indicates that all eight have the same sign. That $\mathcal{R}$ is not sign stable is substantive and is discussed in Appendix J.3. Full sets of marginal contributions for every block are in Appendix Table A24.
Source: Author's calculations using the CEEDD.

Table A24: Marginal contributions by elimination order

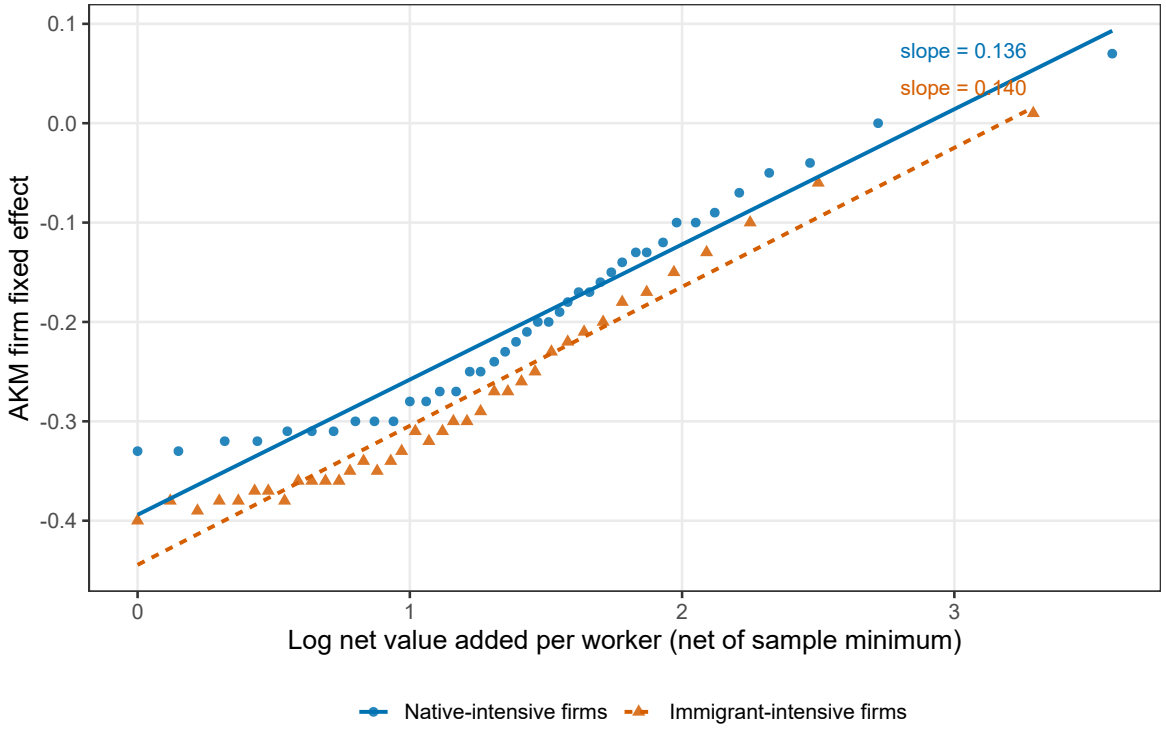
Blocks already eliminated	Gap before (%)	Marginal contribution (pp)
Panel A: Employer substitutability $\mathcal{S}$		
None	14.99	−13.10
$\mathcal{M}$	14.34	−16.74
$\mathcal{V}$	13.58	−13.16
$\mathcal{R}$	9.19	−8.54
$\mathcal{M}, \mathcal{V}$	12.98	−16.98
$\mathcal{M}, \mathcal{R}$	9.13	−8.50
$\mathcal{V}, \mathcal{R}$	8.09	−8.34
$\mathcal{M}, \mathcal{V}, \mathcal{R}$	8.04	−8.31
Panel B: Market structure $\mathcal{M}$		
None	14.99	−0.66
$\mathcal{S}$	1.89	−4.29
$\mathcal{V}$	13.58	−0.61
$\mathcal{R}$	9.19	−0.05
$\mathcal{S}, \mathcal{V}$	0.43	−4.43
$\mathcal{S}, \mathcal{R}$	0.65	−0.02
$\mathcal{V}, \mathcal{R}$	8.09	−0.05
$\mathcal{S}, \mathcal{V}, \mathcal{R}$	−0.25	−0.02
Panel C: Value of non-employment $\mathcal{V}$		
None	14.99	−1.41
$\mathcal{S}$	1.89	−1.46
$\mathcal{M}$	14.34	−1.36
$\mathcal{R}$	9.19	−1.09
$\mathcal{S}, \mathcal{M}$	−2.40	−1.60
$\mathcal{S}, \mathcal{R}$	0.65	−0.90
$\mathcal{M}, \mathcal{R}$	9.13	−1.09
$\mathcal{S}, \mathcal{M}, \mathcal{R}$	0.63	−0.90
Panel D: Skill mix and Technology $\mathcal{R}$		
None	14.99	−5.80
$\mathcal{S}$	1.89	−1.24
$\mathcal{M}$	14.34	−5.20
$\mathcal{V}$	13.58	−5.49
$\mathcal{S}, \mathcal{M}$	−2.40	+3.04
$\mathcal{S}, \mathcal{V}$	0.43	−0.68
$\mathcal{M}, \mathcal{V}$	12.98	−4.94
$\mathcal{S}, \mathcal{M}, \mathcal{V}$	−4.01	+3.74

Notes: All counterfactuals are computed on 2016 data. Each panel reports, for one block b, the quantity $\mathcal{D}(\mathcal{C} \cup \{b\}) - \mathcal{D}(\mathcal{C})$ for each of the eight subsets $\mathcal{C}$ of the other three blocks, where $\mathcal{D}(\cdot)$ is the counterfactual gap of Equation (18). "Gap before" is $\mathcal{D}(\mathcal{C})$. Averaging the eight entries of a panel with the weights of Equation (19) gives that block's Shapley value in Table A23; the "None" row of each panel is the "single block" entry in that table.

Panels are ordered as in Table A17. Reading down a panel shows how the return to one policy lever depends on which others are already in place.

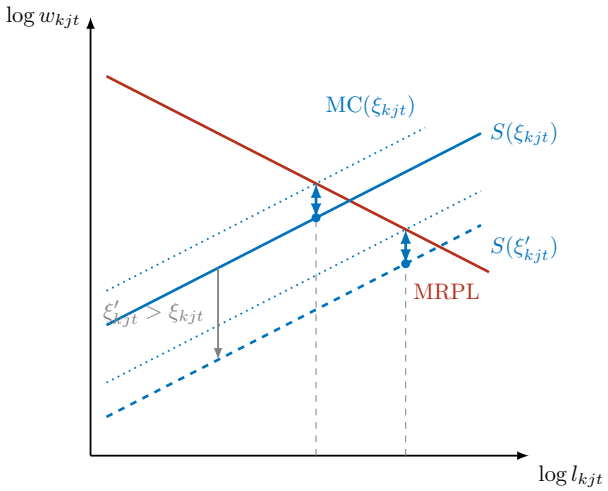
B Appendix Figures

Figure A1: Firm Pay Premia and Net Value Added per Worker, by Immigrant Intensity

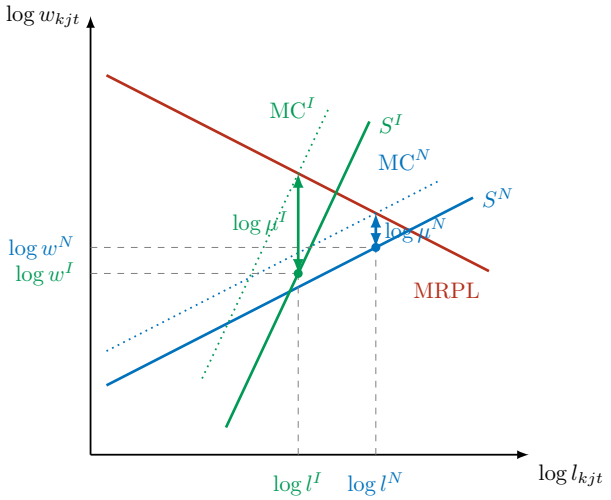


Notes: This figure replicates Figure 3 using net value added per worker in place of revenue per worker. Net value added per worker is computed as the firm's average value added per worker across all available years, winsorized at the 95th percentile with values below the 5th percentile set to zero. The estimated slopes are 0.136 for native-intensive firms and 0.140 for immigrant-intensive firms. Conditional on net value added per worker, immigrant-intensive firms have firm effects that are approximately 4.4 log points lower on average. *Source:* Author's calculations using the CEEDD; value added data come from the NALMF.

Figure A2: Firm-Level Labor Supply, the Supply Shifter, and the Wage Markdown



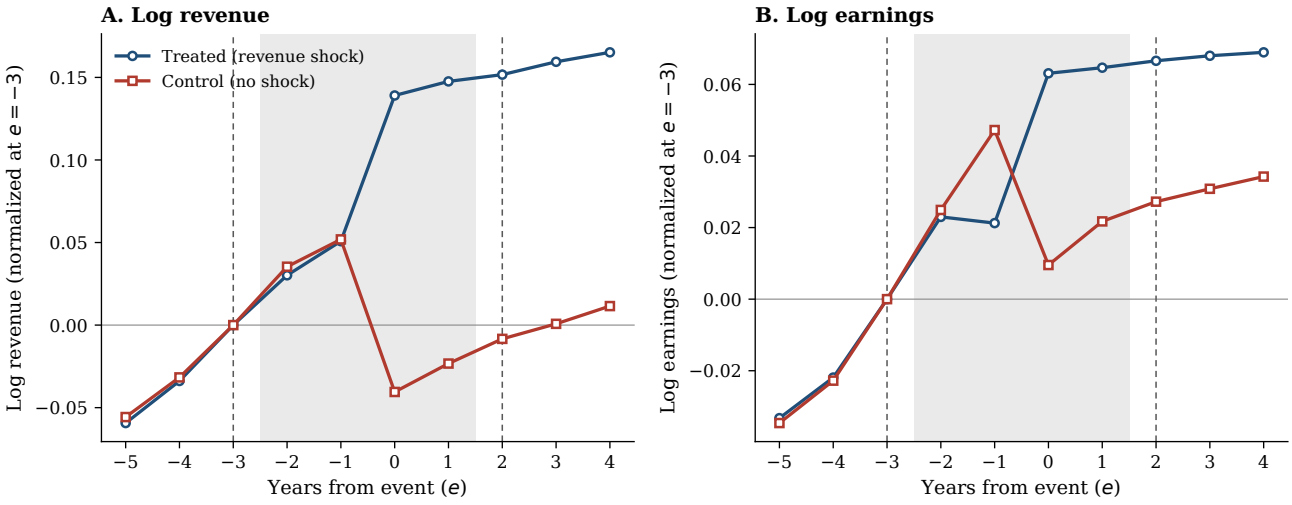
(a) The intercept shifts supply, leaving the markdown unchanged



(b) Less elastic supply implies a larger markdown

Notes: Both panels depict the labor supply curve facing a *single* employer, not the market, and are drawn in logs, so that Equation (4) is a straight line with slope $1/\mathcal{E}_{kjt}$ and intercept $-\xi_{kjt}/\beta_k$. Curves are drawn for a firm small enough that its employment shares can be ignored, in which case $\mathcal{E}_{kjt} = \beta_k \sigma_{kg}$; the share terms in Equation (5) rotate each curve further counterclockwise as the firm's share of type-k employment rises. Solid lines are supply and dotted lines the corresponding marginal cost of labor, which in logs lies a constant $\log(1+1/\mathcal{E}_{kjt})$ above supply. The firm hires where marginal cost equals marginal revenue product and reads the wage off the supply curve; double-headed arrows measure the log markdown of Equation (8), denoted by μ in the figure. Panel (a) varies only ξ_{kjt}, holding β_k and σ_{kg} fixed: a higher intercept shifts supply down in parallel, raising employment and lowering the wage, but leaves the slope and hence the markdown unchanged, as the two arrows are of equal length. Panel (b) holds ξ_{kjt} fixed and lowers $\beta_k \sigma_{kg}$ for immigrants, which steepens supply, widens the gap between supply and marginal cost, and produces a larger markdown, denoted $\log \mu^I$ and $\log \mu^N$. Superscripts I and N denote immigrant and native types. The figure is illustrative and is not drawn to the estimates of Section 5.

Figure A3: Event-Study Representation of the Estimation Procedure



Notes: Mean log revenue (Panel A) and mean log earnings (Panel B) for firms receiving an above-median versus a below-median log revenue change at event time zero, pooling all worker types. Both series are normalized to zero at event time −3, the base period of the long difference. Event years are stacked and each firm is weighted by employment; no controls or fixed effects are applied. The shaded area denotes the event times at which the orthogonality condition need not hold. The vertical dashed lines at event times −3 and +2 mark the endpoints of $\Delta_{e,e'}$ with $e = 2$ and $e' = 3$.

Figure A4: DiD Representation of the Estimation Procedure, by Worker Type: Revenue per Worker

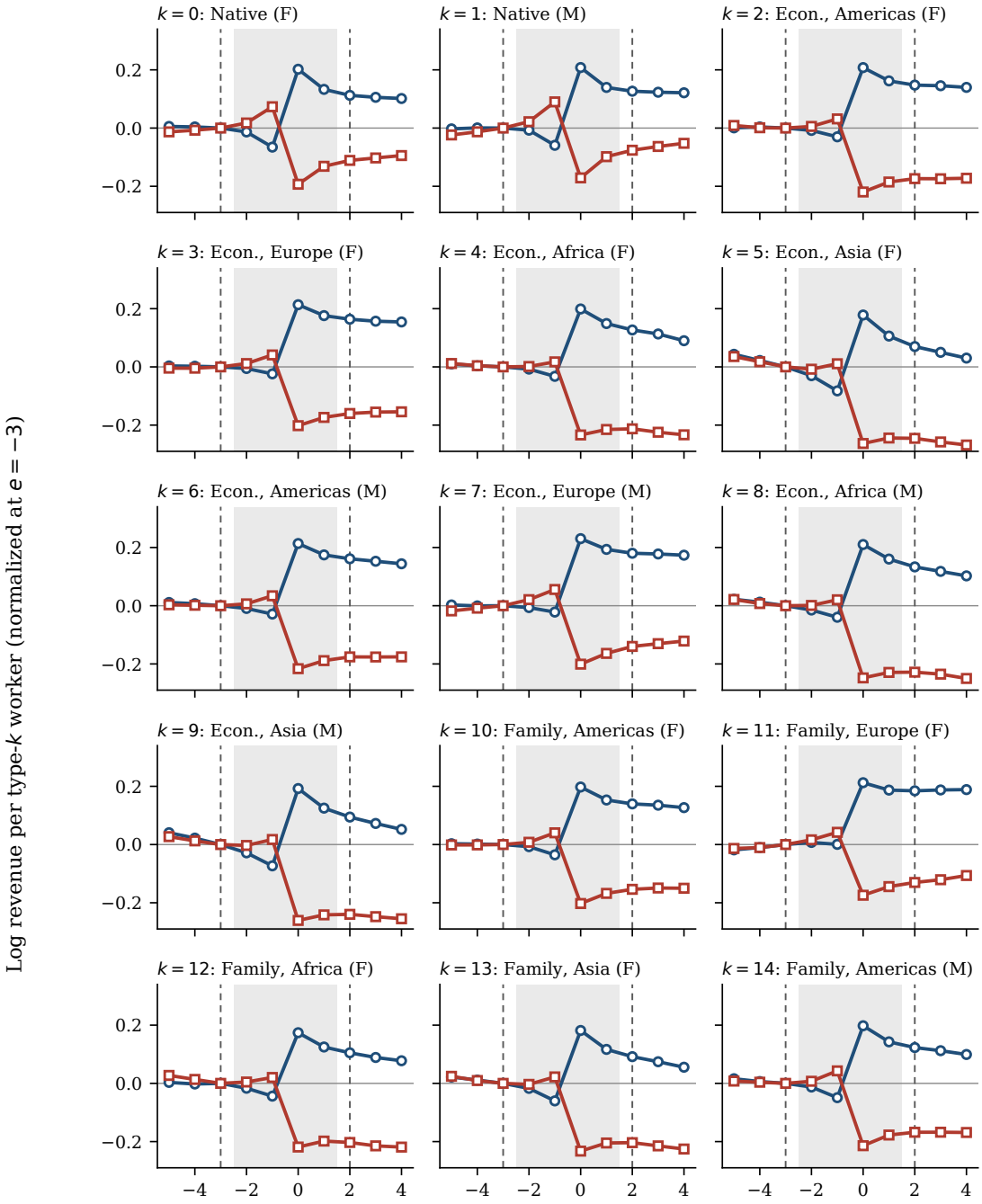
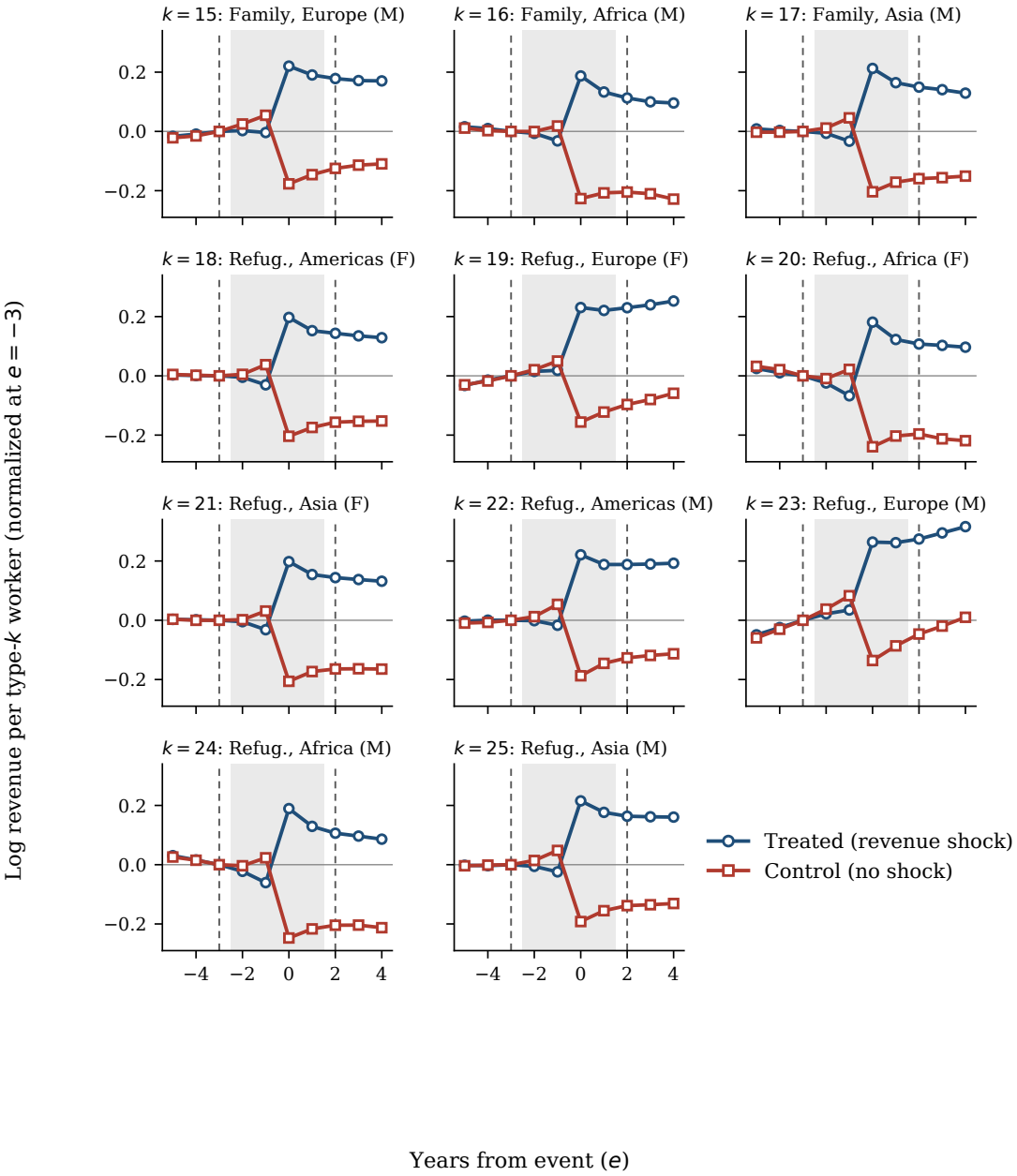


Figure A4: DiD Representation of the Estimation Procedure, by Worker Type: Revenue per Worker *(continued)*



Notes: Mean log revenue per type-k worker for firms receiving an above-median versus a below-median log revenue change at event time zero, reported separately for each worker type. Series are normalized to zero at event time -3 and share a common vertical scale across panels. Construction is otherwise identical to Figure A3. The residual type is omitted. Shaded regions denote event times at which the orthogonality condition need not hold; vertical dashed lines mark the endpoints of $\Delta_{e,e'}$ with $e = 2$ and $e' = 3$. *Source:* Author's calculations using the CEEDD.

Figure A5: DiD Representation of the Estimation Procedure, by Worker Type: Earnings

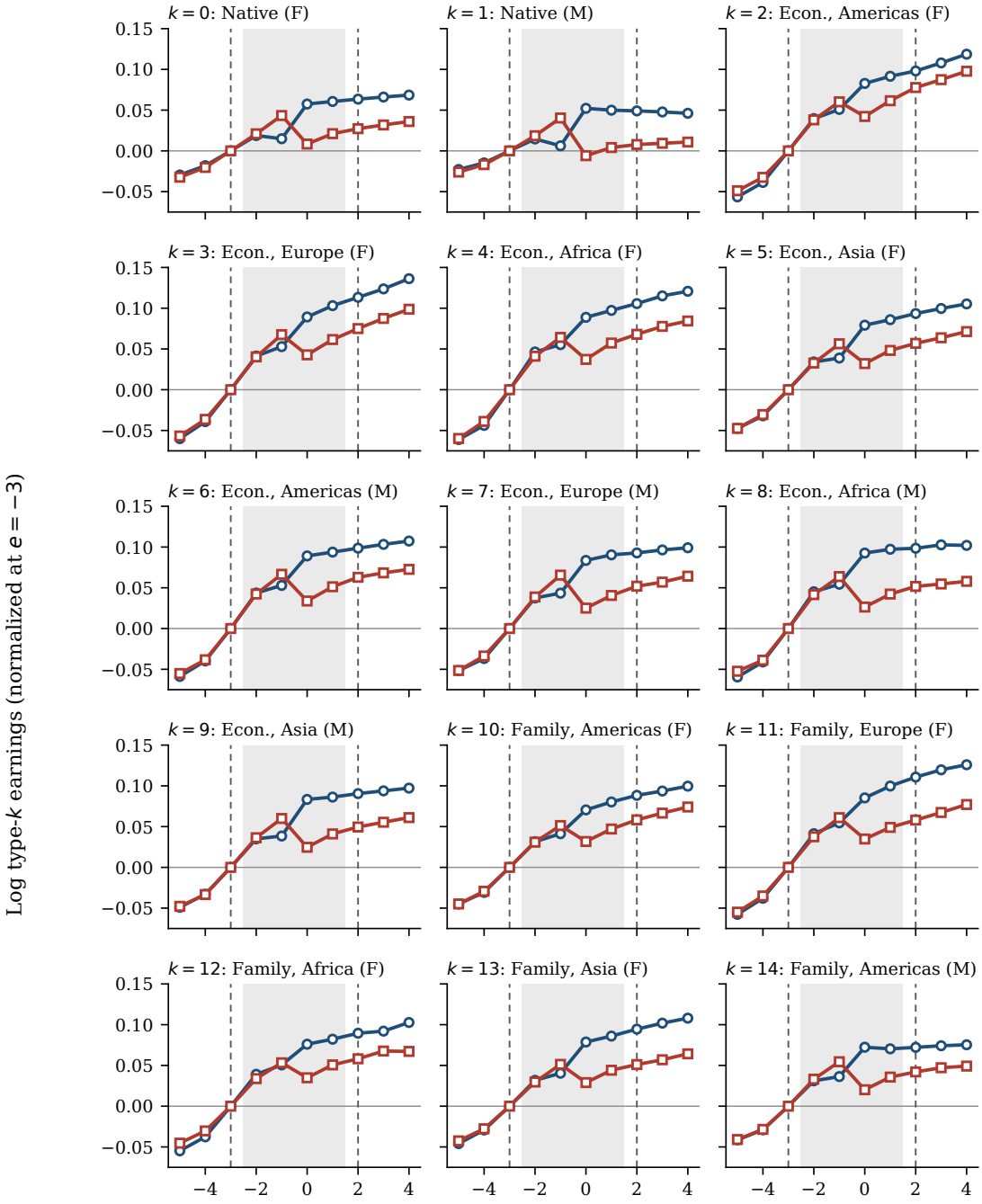
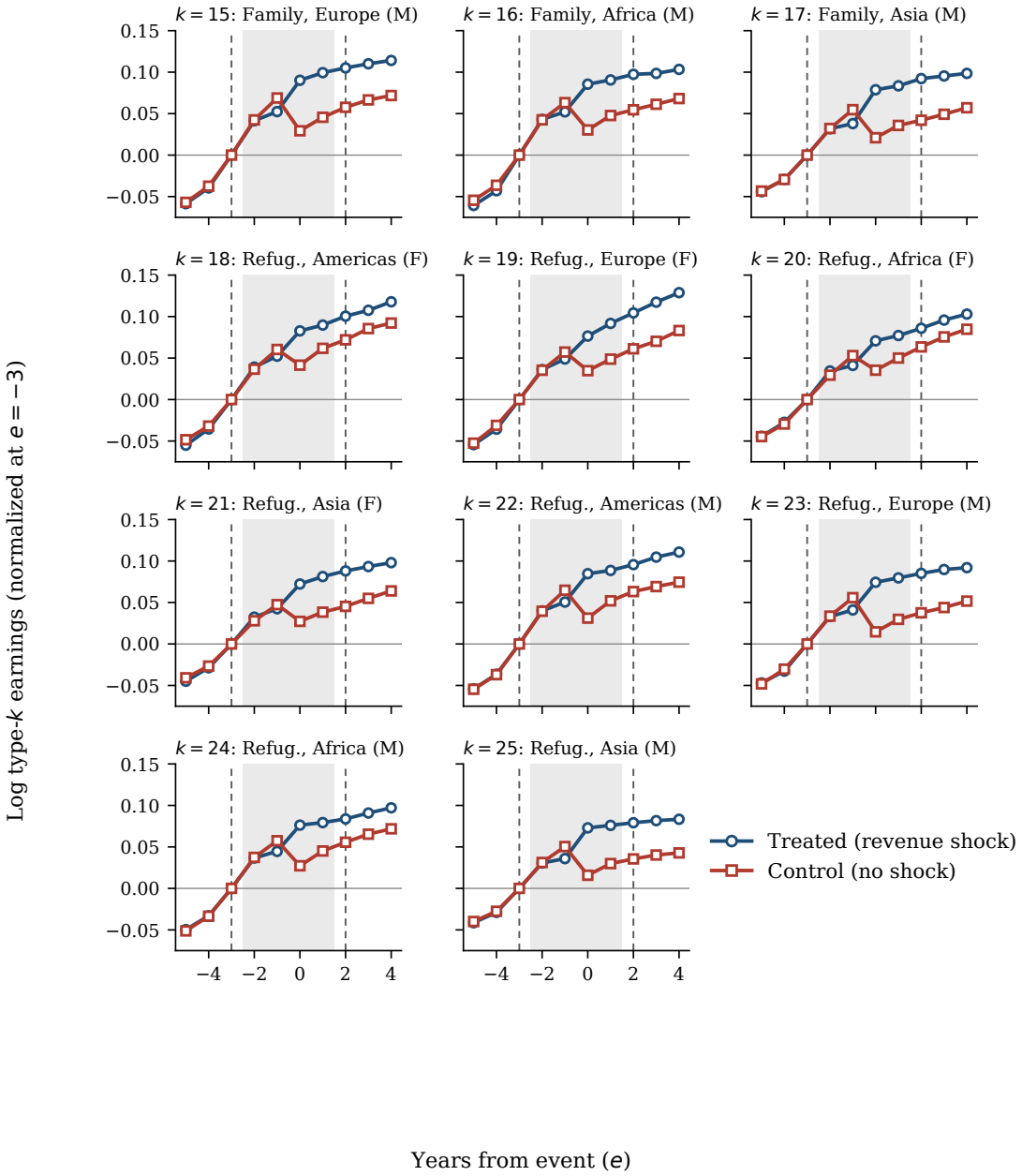
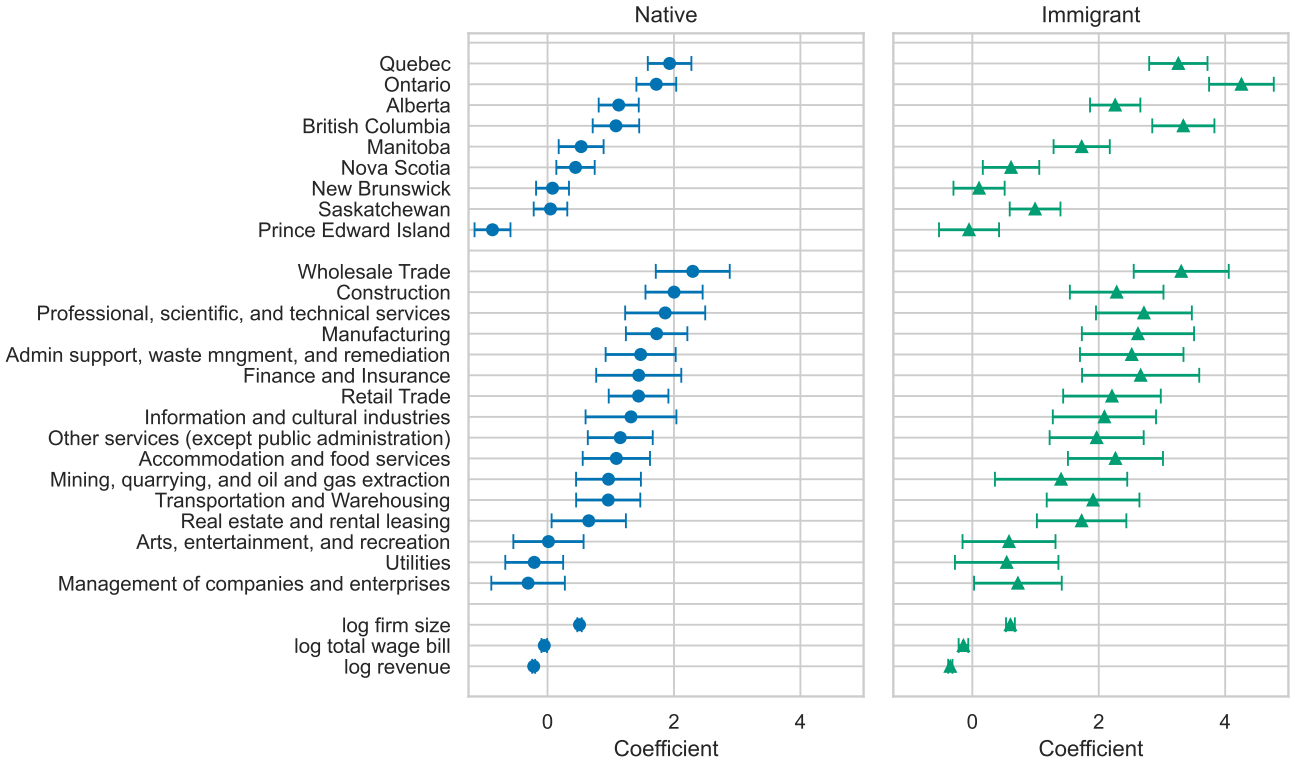


Figure A5: DiD Representation of the Estimation Procedure, by Worker Type: Earnings *(continued)*



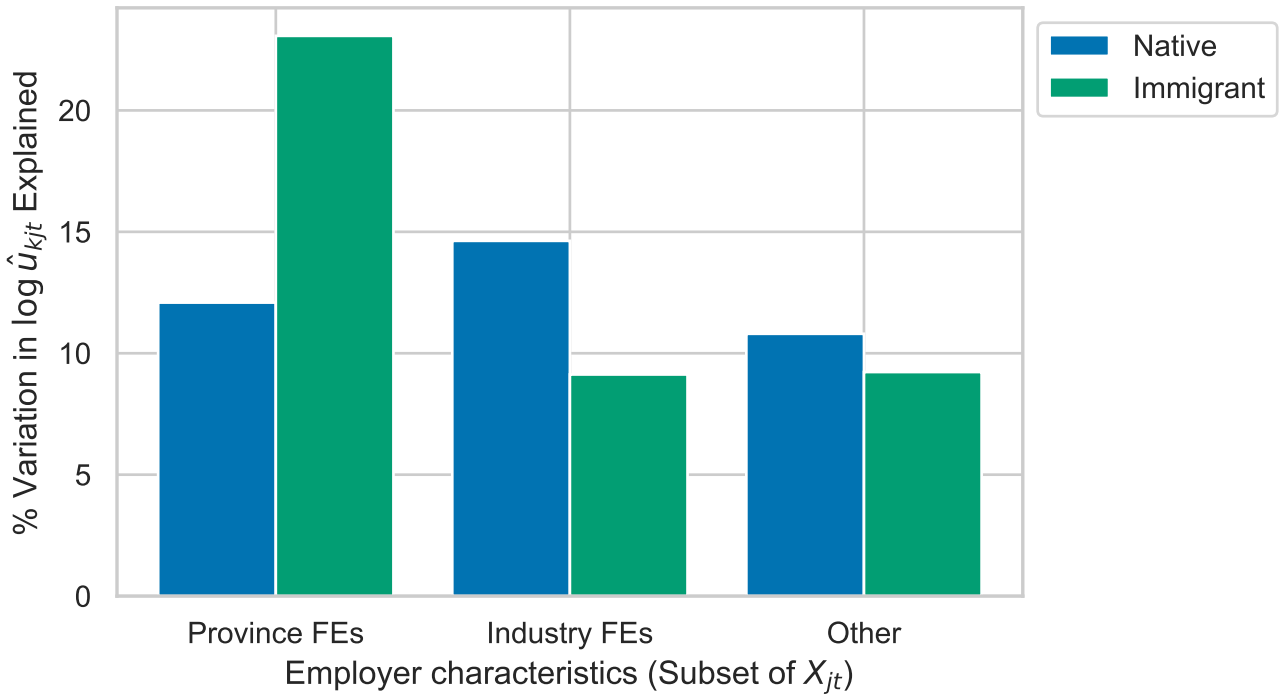
Notes: Mean log type-k earnings for firms receiving an above-median versus a below-median log revenue change at event time zero, reported separately for each worker type. Series are normalized to zero at event time -3 and share a common vertical scale across panels. Construction is otherwise identical to Figure A3. The residual type is omitted. *Source:* Author's calculations using the CEEDD.

Figure A6: Estimated Supply-Shifter Coefficients on Employer Characteristics, by Immigrant Status



Notes: Coefficients and fixed effects from Equation (46). The dependent variable is the recovered supply shifter $\widehat{\log \xi_{kjt}}$ of Equation (45): the part of a group's attachment to an employer not accounted for by the wage or by the employer's share of type-k employment in the market. Bars are 95 percent confidence intervals from standard errors clustered at the firm. *Source:* Author's calculations using the CEEDD.

Figure A7: Incremental R^2 of the Supply-Shifter Regression, by Immigrant Status



Notes: Each panel reports, for one of the three blocks of right-hand-side variables in Equation (46), the drop in R^2 from removing that block from the full specification, $\Delta R^2 \equiv R^2_{(1)} - R^2_{(2)}$, estimated separately for each of the 26 worker types and then aggregated to immigrant status. The blocks are province effects, two-digit NAICS industry effects, and the observable employer characteristics in X_{jt} (revenue, headcount, total wage bill). The dependent variable is the recovered supply shifter $\widehat{\log \xi_{kjt}}$ of Equation (45). Averaging across types, the location increment is 0.219 for immigrant types and 0.126 for native types, the industry increment 0.094 and 0.139, and the employer-characteristics increment 0.097 and 0.107. Because the estimated fixed effects are normalized against an omitted category, their levels are not comparable across specifications, which is why explanatory power rather than coefficients is compared here. *Source:* Author's calculations using the CEEDD.

C Immigration in Canada

Foreign-born individuals living in Canada may hold temporary resident status, permanent resident status, or citizenship. Permanent residents have the right to live, work, and study in Canada indefinitely; the only difference between permanent residency and citizenship is a limited set of additional rights, including eligibility to run for political office and hold certain government jobs requiring high-level security clearance. This is why I do not distinguish immigrants who later obtained citizenship from those who remain permanent residents. Individuals are granted permanent residence through one of three admission classes: the *Economic Class*, the *Family Class*, or the *Refugee and Humanitarian Class*.

C.1 The Economic Class

The *Economic Class* grants permanent residence to individuals primarily on the basis of their human capital characteristics. This class encompasses a range of federal and subnational programs designed to select applicants with skills and experience expected to contribute to the Canadian labor market. Because admission is conditioned on education, experience, and language, this group's credentials are the most directly comparable to natives', which makes it the natural place to look when measured skill fails to account for the pay gap (Section 3). There are three main federal programs within the *Economic Class*: The *Federal Skilled Workers Program (FSWP)*, the *Canadian Experience Class (CEC)*, and the *Federal Skilled Trades Program (FSTP)*. The FSWP is a points-based system that was first introduced in 1967. Applicants to the FSWP are assigned points based on education, age, skilled work experience, proficiency in English or French, ability to integrate into Canada, and whether the applicant has a job offer from a Canadian employer. The federal government sets a minimum points threshold, and applicants who meet or exceed this threshold satisfy the program's selection criteria. In contrast to the FSWP, the CEC and FSTP evaluate applicants on a pass/fail basis. The CEC was introduced in 2008 and requires applicants to meet a minimum Canadian work experience requirement (initially two years, reduced to one year in 2012), while the FSTP was introduced in 2013 and requires applicants to have a Canadian certificate of qualification in a skilled trade or a job offer from a Canadian employer.[1]

At the subnational level, the *Provincial Nominee Program (PNP)* allows provinces and territories to design immigration streams based on local economic and demographic needs. Manitoba signed the first PNP agreement in 1996, and since then, all provinces and territories except Quebec and Nunavut have joined the program. The immigration streams in the PNP are diverse, with over 50

1. Since 2015, the FSWP, FSTP, and CEC have been centrally managed by the federal *Express Entry (EE)* system. The EE system is an application management system, not a stand-alone program. In the EE system, applicants must first meet the minimum eligibility requirements of the FSWP, FSTP, or CEC to enter the EE pool, where they subsequently receive a Comprehensive Ranking System (CRS) score. Applicants in the EE pool with CRS scores above a minimum threshold are invited to apply for permanent residency.

distinct PNP streams across 11 participating provinces and territories in 2011 (Citizenship and Immigration Canada 2011). PNP streams commonly focus on selecting immigrants who work in high-demand occupations or have a job offer from an employer within the province. If the province or territory determines that an applicant meets the eligibility criteria of a particular stream, it formally selects the applicant for permanent residence through a process called a provincial nomination. Once nominated, the applicant receives the *Provincial Nominee Class (PNC)* designation and the federal government subsequently conducts medical, security, and criminal background checks and confirms that the applicant intends to reside in the nominating province before granting permanent residence.

The province of Quebec has its own program within the *Economic Class* that is distinct from the FSWP, CEC, FSTP, and PNP, introduced in 1991 and known as *Programme régulier des travailleurs qualifiés (PRTQ)* (replaced by *Programme de sélection des travailleurs qualifiés (PSTQ)* in 2024). This program evaluates applicants using a points-based system similar to the FSWP, but includes additional criteria for French proficiency and family connections in Quebec.

C.2 The Family Class

The *Family Class* was introduced in 1976 as an admission category for applicants who have family members who are Canadian citizens or permanent residents. Through the *Family Class*, permanent residents or citizens of Canada can sponsor other family members, including spouses, dependents, and parents, and must commit to financially supporting sponsored relatives for several years. For applicants who intend to reside in Quebec, the process includes an additional step in which the Quebec government must approve the sponsor.

C.3 The Refugee and Humanitarian Class

The *Refugee and Humanitarian Class* grants permanent residence to individuals based on need for protection. It was formally introduced with the *Immigration Act* of 1976, which made refugee protection a permanent part of Canada's immigration system. Individuals are selected on humanitarian grounds when they face serious risks such as persecution, violence, or hardship in their home country, or when exceptional compassionate circumstances apply. Refugees receive additional support from the federal government compared to other immigrants when they first arrive in Canada, including income assistance, housing help, and settlement services.

C.4 Temporary residents

The three admission classes discussed above are only available for permanent residents. Foreign-born individuals who are living or working in Canada and do not have citizenship or permanent

residency status must have temporary resident status. Temporary residents are given the right to live, work, or study in Canada for a limited time only and include international students, temporary foreign workers (TFWs) with employer-specific visas admitted under the *Temporary Foreign Worker Program (TFWP)*, and other TFWs with open work permits admitted through the *International Mobility Program (IMP)*. Due to the unique legal restrictions associated with temporary visas, I exclude temporary residents from my analysis. For an analysis of earnings inequality between TFWs and permanent residents, see Kroft et al. (2025).

D Estimating firm pay premiums using the AKM model

D.1 Sample details

The model is estimated on the AKM sample described in Section 2, with one additional restriction. Firm effects are only identified for firms in a "connected set," meaning that there exists a worker who moves between them at some point during the sample period. Thus, an important step prior to estimating the AKM-style model involves restricting the sample to the largest connected set of workers and firms.[2] Thus, firms that are not in the connected set are excluded (fewer than 5%).

D.2 AKM estimation details

The AKM model (Abowd, Kramarz, and Margolis 1999) assumes that the log of an individual's earnings can be decomposed into the sum of a person effect, a firm effect, a time-varying index of individual characteristics, and a residual. Formally, the AKM model posits that the log earnings of individual i at firm j at time t can be written as:

$$y_{it} = \alpha_i + \psi_{\mathbf{j}(i,t)} + X'_{it}\beta + \varepsilon_{it}, \tag{20}$$

where α_i is the person effect for individual i, ψ_j is the firm effect for firm j, X_{it} is a vector of time-varying characteristics for individual i at time t, β is a conformable vector of coefficients, and ε_{it} is the residual. Included in the vector X_{it} are variables for marital status, province of residence, year effects, and controls for age.

To control for age effects, I include a quartic polynomial in normalized age in the vector of controls X_{it}. Since I also include year effects in X_{it}, the linear term of the polynomial in age is not identified.[3] Therefore, I omit the linear term of the polynomial in age and also normalize age by subtracting and dividing by the age at which the earnings profile is at a maximum. As explained by Card et al. (2018), omitting the linear term of the polynomial without including the additional normalization of age can bias the person effects upward.

For the estimator of the firm effects to be unbiased, firm-to-firm mobility must be uncorrelated with time-varying unobservables. This is often referred to as the "exogenous mobility assumption." Dostie et al. (2023) provide evidence supporting the exogenous mobility assumption in the CEEDD from 2001 to 2013, a period similar to the main sample (see Section 2).

2. To extract the largest connected set of workers and firms, I use the igraph package. The employer-employee data can be viewed as a graph where the firms are nodes and the edges are worker flows between firms. The largest connected set is equivalent to the maximal connected component of the worker-firm graph.

3. Age is a linear function of year and birth year, and the person effects are collinear with birth year.

D.3 What the AKM firm effect can and cannot capture

This appendix formalizes the claim in Section 3 that the immigrant–native gap in AKM firm premia is a lower bound on the contribution of employers, and identifies which component of wage-setting power it misses.

The wage equation in logs. Taking logs of Equation (7),

$$\log w_{kjt} = \log \mathrm{MRPL}_{kjt} + \log\left(\frac{\mathcal{E}_{kjt}}{1+\mathcal{E}_{kjt}}\right). \tag{21}$$

Substituting Equation (5) and adding and subtracting the log pass-through evaluated at zero shares splits the second term into

$$\log\left(\frac{\mathcal{E}_{kjt}}{1+\mathcal{E}_{kjt}}\right) = \underbrace{\log\left(\frac{\beta_k\sigma_{kg}}{1+\beta_k\sigma_{kg}}\right)}_{\equiv\, m_{kg}} + \underbrace{\log\left(\frac{\mathcal{E}_{kjt}}{1+\mathcal{E}_{kjt}}\right) - \log\left(\frac{\beta_k\sigma_{kg}}{1+\beta_k\sigma_{kg}}\right)}_{\equiv\, d_{kjt}}. \tag{22}$$

The first term, m_{kg}, is equivalent to one minus the markdown, and it is the pass-through a type-k worker faces at an employer with negligible share, and is therefore common to every firm in market g. The second, $d_{kjt} \leq 0$, is the additional pass-through generated by firm j's share of type-k employment, and is zero for an atomistic employer. The split is exactly the two sources of wage-setting power: m_{kg} is the substitutability of employers and d_{kjt} is concentration.

What movers identify. The AKM specification is Equation (20), with ψ_j identified from within-worker wage changes at job switches within a connected set. Consider a type-k worker moving from j to j' within market g between t and $t+1$. From Equations (21) and (22),

$$\Delta \log w = \big(d_{kj't+1} - d_{kjt}\big) + \big(\log \mathrm{MRPL}_{kj't+1} - \log \mathrm{MRPL}_{kjt}\big), \tag{23}$$

in which m_{kg} has differenced out. Because a worker's type is fixed, m_{kg} enters every one of her wage observations identically and is absorbed into $\hat{\alpha}_i$, where it is observationally equivalent to human capital. Two workers with identical marginal revenue products at identical employers, differing only in $\beta_k\sigma_{kg}$, generate identical $\hat{\psi}_j$ and different $\hat{\alpha}_i$.

The component that survives. Equation (23) shows that AKM is not blind to wage-setting power. What differences out is the *level* of the markdown, common to all employers a worker could move to; what survives is its *dispersion* across employers.

Qualifications. First, d_{kjt} depends on the firm–type pair, not the firm alone, so the markdown component AKM does capture enters as a type-specific firm effect rather than the additive ψ_j

that AKM imposes; the resulting $\hat{\psi}_j$ is an employment-weighted average across types. This is a further reason the premium gap understates employer-driven differences, and a motivation for the interacted specifications in Card, Cardoso, and Kline (2016). Second, the argument conditions on within-market mobility. Moves across markets identify differences in m_{kg} across g for a given type, but not the level of m_{kg}, which is what enters the pay gap. Third, Equation (23) inherits the exogenous mobility assumption of Abowd, Kramarz, and Margolis (1999): if movers select on d or on match productivity, $\hat{\psi}_j$ is contaminated regardless.

E Concentration measures: derivations

This appendix collects the algebra behind the concentration measures in Section 3. Notation follows the main text: n_{kj} is the number of type-k workers at firm j, $N_{kg} = \sum_{j \in \mathcal{J}_g} n_{kj}$, $N_k = \sum_g N_{kg}$, and $s_{kg} = N_{kg}/N_k$.

E.1 The HHI as a with-replacement collision probability

Fix any population of N type-k workers distributed across firms with counts n_j, and let $H = \sum_j (n_j/N)^2$ and $C = \sum_j n_j(n_j-1)/[N(N-1)]$. Drawing two workers with replacement, the pair coincides at firm j with probability $(n_j/N)^2$, so H is the with-replacement collision probability; drawing without replacement, the second worker is at j with probability $(n_j-1)/(N-1)$, giving C. Then

$$C \;=\; \sum_j \frac{n_j^2 - n_j}{N(N-1)} \;=\; \frac{1}{N-1}\left(\sum_j \frac{n_j^2}{N} - \frac{\sum_j n_j}{N}\right) \;=\; \frac{NH-1}{N-1},$$

using $\sum_j n_j = N$, and rearranging,

$$H \;=\; C + \frac{1-C}{N}. \tag{24}$$

The identity holds at any level of aggregation. Applied within a market it gives $H_{kg} = C_{kg} + (1 - C_{kg})/N_{kg}$, the mechanical inflation discussed in the main text: holding the true collision probability fixed, the group with fewer workers in the cell records the higher HHI. The bias vanishes as $N \to \infty$, which is why the two measures agree to reported precision in the CEEDD at the aggregate level, but it is not innocuous in comparisons across groups of unequal size or across finely cut markets in which one group is thin.

E.2 Exact within–between decomposition of the collision probability

Two type-k workers collide at a firm if and only if they are in the same market and collide within it. Formally, split Equation (2) by market and multiply and divide by $N_{kg}(N_{kg}-1)$:

$$C_k \;=\; \sum_{g \in \mathcal{G}} \sum_{j \in \mathcal{J}_g} \frac{n_{kj}(n_{kj}-1)}{N_k(N_k-1)} \;=\; \sum_{g \in \mathcal{G}} \frac{N_{kg}(N_{kg}-1)}{N_k(N_k-1)} \underbrace{\sum_{j \in \mathcal{J}_g} \frac{n_{kj}(n_{kj}-1)}{N_{kg}(N_{kg}-1)}}_{=C_{kg}}.$$

With $C_k^B \equiv \sum_g N_{kg}(N_{kg}-1)/[N_k(N_k-1)]$ and weights $\pi_{kg} \equiv N_{kg}(N_{kg}-1)/[N_k(N_k-1)C_k^B]$, which sum to one by construction, this is

$$C_k \;=\; C_k^B \times \sum_g \pi_{kg} C_{kg} \;\equiv\; C_k^B \times C_k^W,$$

which is Equation (3). The decomposition is exact and both factors are probabilities of the same kind. The weights π_{kg} are the probabilities that a colliding-in-market pair is in market g, so C_k^W is itself interpretable as the collision probability conditional on the pair sharing a market.

E.3 Decomposition of the overall HHI

The overall HHI $H_k \equiv \sum_{j\in\mathcal{J}}(n_{kj}/N_k)^2$ also splits across the partition, because the with-replacement draw factors in the same way. Writing $H_{kg} = \sum_{j\in\mathcal{J}_g}(n_{kj}/N_{kg})^2$,

$$H_k \;=\; \sum_g s_{kg}^2 H_{kg} \;=\; H_k^B \times \sum_g \omega_{kg} H_{kg}, \qquad H_k^B \equiv \sum_g s_{kg}^2, \quad \omega_{kg} \equiv \frac{s_{kg}^2}{H_k^B},$$

which is exact. The difficulty is interpretive: the weights ω_{kg} are squared employment shares renormalized, and the resulting average does not correspond to the concentration faced by any worker, nor to a conditional probability of the kind C_k^W delivers. The quantity reported in the HHI panel of Table 3 is instead the conventional employment-weighted average,

$$H_k^W \;\equiv\; \sum_g s_{kg} H_{kg},$$

the average within-market concentration faced by a randomly chosen type-k worker, and consequently $H_k^W \times H_k^B$ reproduces H_k only approximately. The approximation is close for immigrants, whose employment is concentrated in relatively few markets so that the two weighting schemes nearly coincide ($0.0482 \times 0.0209 = 0.0010$, matching H_k), and poorer for natives, whose employment is spread across many markets ($0.0643 \times 0.0070 = 0.0005$ against $H_k = 0.0008$). The trade-off between an exact decomposition with uninterpretable weights and an interpretable average that does not multiply back is the reason the main text leads with the collision probability, which faces neither problem. The between-market terms are nearly identical across the two panels of Table 3 because C_k^B and H_k^B differ only by the correction in Equation (24), applied at the market rather than the firm level, where it is negligible.

F Derivation and Interpretation of the Labor Supply Curve

Throughout this appendix we fix a worker type k and a period t and suppress both subscripts. They are restored where the results are stated.

F.1 Derivation of Lemma 4.1

Preferences. A worker i in market g values employer $j \in \mathcal{J}_g$ at

$$v_{ij} = \delta_j + \epsilon_{ij}, \qquad \delta_j \equiv \beta \log w_j + \xi_j, \tag{25}$$

where $j = 0$ denotes non-employment, w_0 is its value, and $\xi_0 = 0$. The term $\beta \log w_j$ is the payoff from pay. The term ξ_j collects everything else that is common to workers of the type: commuting distance, scheduling, workplace language, injury risk, how welcoming the employer is to a given group, and how well known it is to that group. The shock ϵ_{ij} is the only component along which two workers of the same type, facing the same wages, differ.

Distributional assumption. The vector $(\epsilon_{i0}, \{\epsilon_{ij}\}_{j \in \mathcal{J}_g, g})$ is drawn i.i.d. across workers from the generalized extreme value distribution

$$F(\epsilon) = \exp\left[- e^{-\epsilon_0} - \sum_g \Big(\sum_{j \in \mathcal{J}_g} e^{-\sigma_g \epsilon_j} \Big)^{1/\sigma_g} \right], \qquad \sigma_g \geq 1. \tag{26}$$

Each local labor market is a nest, and non-employment is its own nest. The nesting parameter is $\tilde{\sigma}_g = 1 - 1/\sigma_g \in [0, 1)$. When $\sigma_g = 1$, the shocks are independent type I extreme value and the model reduces to multinomial logit. As σ_g rises, shocks within market g become more correlated, and employers in the same market become closer substitutes. The restriction $\sigma_g \geq 1$ ensures consistency with random utility maximization (McFadden 1978).

Choice probabilities. Let $D_g \equiv \sum_{l \in \mathcal{J}_g} e^{\sigma_g \delta_l}$. Under (26), the probability that a worker chooses $j \in \mathcal{J}_g$ is

$$s_j = \underbrace{\frac{e^{\sigma_g \delta_j}}{D_g}}_{s_{j|g}} \times \underbrace{\frac{D_g^{1/\sigma_g}}{e^{\delta_0} + \sum_{g'} D_{g'}^{1/\sigma_{g'}}}}_{s_g}, \qquad s_0 = \frac{e^{\delta_0}}{e^{\delta_0} + \sum_{g'} D_{g'}^{1/\sigma_{g'}}}, \tag{27}$$

where $s_{j|g}$ is the share of j within market g and s_g is the share of market g.

Quasi-supply. Dividing s_j by s_0 gives $\log(s_j/s_0) = \sigma_g\delta_j - \delta_0 + (1/\sigma_g - 1)\log D_g$. Substituting $\log D_g = \sigma_g\delta_j - \log s_{j|g}$ from the definition of $s_{j|g}$, and using $\xi_0 = 0$, yields

$$\log\frac{s_j}{s_0} = \beta\log\frac{w_j}{w_0} + \tilde{\sigma}_g\log s_{j|g} + \xi_j. \tag{28}$$

With a continuum of workers of each type, choice probabilities equal employment shares. Equation (28) therefore holds at any wage vector, including the equilibrium one. Restoring the subscripts k and t and exponentiating gives Equation (4).

F.2 An observationally equivalent information-based reading

The derivation above attributes dispersion in choices to idiosyncratic tastes. I now show that the same choice rule, and the same expected-utility function, arise in a model with no taste shocks, in which workers are imperfectly informed about pay and attention is costly.

The attention problem. Suppose workers do not observe posted wages without effort. For a given realization of wages, a worker chooses the probabilities p_j with which she ends up at each option. Doing so requires information, and information is costly. She solves

$$V = \max_{\{p_j\}} \sum_j p_j\big(\beta\log w_j + \log q_j\big) \;-\; C(p) \quad \text{subject to} \quad \sum_j p_j = 1, \tag{29}$$

where q_j is her default propensity to consider employer j absent any information, and

$$C(p) = \sum_g \Big[\tfrac{1}{\sigma_g}\sum_{j\in\mathcal{J}_g} p_j\log p_j + \big(1-\tfrac{1}{\sigma_g}\big)\,p_g\log p_g\Big], \qquad p_g \equiv \sum_{j\in\mathcal{J}_g} p_j. \tag{30}$$

The coefficient β is the weight on pay relative to the unit cost of information. When $\sigma_g = 1$, the objective in Equation (29) equals $\sum_j p_j\beta\log w_j - \sum_j p_j\log(p_j/q_j)$. This is the rational inattention problem of Matějka and McKay (2015), in which the cost is the relative entropy of the chosen probabilities to the default. For $\sigma_g > 1$, Equation (30) belongs to the class of Bregman information costs of Fosgerau et al. (2020). Below, I show that this cost term $C(p)$ can be decomposed into two terms: the first is the cost of learning which nest to enter, and the second is the cost of learning which employer to choose within a nest.

Solution. Let $\delta_j \equiv \beta\log w_j + \log q_j$. Substitute the adding-up constraint, $p_0 = 1 - \sum_{j\neq 0} p_j$, and the definition $p_g = \sum_{i\in\mathcal{J}_g} p_i$ directly into the objective, and choose $\{p_j\}_{j\neq 0}$. Non-employment is alone in its nest, so its cost term reduces to $p_0\log p_0$. For $j \in \mathcal{J}_g$, we have $\partial p_g/\partial p_j = 1$ and

$\partial p_0/\partial p_j = -1$. The first-order condition is therefore

$$\delta_j - \delta_0 = \tfrac{1}{\sigma_g}\big(\log p_j + 1\big) + \big(1 - \tfrac{1}{\sigma_g}\big)\big(\log p_g + 1\big) - \big(\log p_0 + 1\big) = \tfrac{1}{\sigma_g}\log p_j + \big(1 - \tfrac{1}{\sigma_g}\big)\log p_g - \log p_0,$$

where the constants cancel. Writing $p_j = p_{j|g}\, p_g$, this becomes $\delta_j - \delta_0 = \frac{1}{\sigma_g}\log p_{j|g} + \log p_g - \log p_0$. Adding $\big(1 - \frac{1}{\sigma_g}\big)\log p_{j|g}$ to both sides,

$$\log \frac{p_j}{p_0} = \log p_{j|g} + \log p_g - \log p_0 = \beta \log \frac{w_j}{w_0} + \tilde{\sigma}_g \log p_{j|g} + \log \frac{q_j}{q_0}, \qquad \tilde{\sigma}_g = 1 - \tfrac{1}{\sigma_g}. \tag{31}$$

This is Equation (28), with

$$\xi_j = \log \frac{q_j}{q_0}. \tag{32}$$

Implications. Under the information reading, each parameter has a counterpart in the cost of attention. Using $p_j = p_{j|g}\, p_g$, the cost in Equation (30) can be written as

$$C(p) = \sum_g \Big[p_g \log p_g + \tfrac{1}{\sigma_g}\, p_g \sum_{j \in \mathcal{J}_g} p_{j|g} \log p_{j|g} \Big]. \tag{33}$$

The first term is the cost of learning which nest to enter. The second is the cost of learning which employer to choose within a nest, and it is scaled by $1/\sigma_g$.

The three objects are then read as follows. First, ξ_{kjt} measures how salient employer j is to type-k workers, not how much they like it. Second, a low β_k reflects costly attention to pay, not indifference to it. Third, σ_{kg} measures how cheaply type-k workers can compare employers within nest g. A high σ_{kg} means that, once a worker attends to a nest, telling its employers apart is inexpensive. A low σ_{kg} means that workers may know a local labor market well enough to search in it, yet find it costly to learn which of its employers pay more. Under the taste reading, the same parameter instead reflects how strongly a worker's tastes are correlated across employers in the same nest.

Both readings imply the same supply curve, the same elasticity in Equation (5), and the same markdown in Equation (8). Nothing in the estimation or in the counterfactuals of Section 6 distinguishes them.

The information reading is a natural one for the immigrant–native pay gap. Recent immigrants arrive with little knowledge of the host-country labor market. They know fewer employers, they know less about what those employers pay, and they often learn about openings through narrow, group-specific channels (Munshi 2003; Dustmann et al. 2016). Under this reading, employers have wage-setting power over immigrants because immigrants are less informed about their alternatives, not because they care less about pay.

The distinction matters only for how the parameters are interpreted. A steeper immigrant supply curve is consistent with immigrants placing more weight on non-wage attributes of employers, with immigrants knowing less about which local employers are hiring and at what pay, or with both. I therefore report (β_k, σ_{kg}) and the markdowns they imply, while remaining agnostic about the specific source of the frictions generating β_k and σ_{kg}.

What the frictions share is more important than what separates them. Under either reading, immigrants effectively face fewer outside options than natives with the same skills. A given wage cut therefore causes fewer of them to leave, and the labor supply curve an employer faces for immigrants is steeper than the one it faces for natives. An employer with wage-setting power responds by marking down immigrant pay by more, relative to the marginal revenue product of labor. This is the mechanism in Panel (b) of Figure A2. Two workers who are equally productive at the same firm can then be paid differently, and the difference reflects their outside options and employer wage-setting power rather than worker productivity. It is this common implication, rather than any particular microfoundation, that the paper takes to the data.

Identification. The identification argument of Section 4.5 uses only Equation (28) and a restriction on the timing of changes in ξ_{kjt} (Assumption 4.4). Under the information reading, $\xi_{kjt} = \log(q_{kjt}/q_{k0t})$, and the assumption becomes a restriction on the timing of changes in how visible employer j is to type-k workers. Persistent differences in visibility are permitted. An employer may be well known among a particular immigrant group because it is co-ethnically owned, located in an enclave, or long established in the community. All such differences are absorbed by the fixed component $\bar{\xi}_{kj}$ and differenced out. What must be ruled out is that an employer becomes durably more visible to a group in the same years that its productivity changes.

Referral networks are the natural way this could happen, and they fit the information reading especially well. Suppose a firm expands after a productivity shock and hires immigrant workers, and that those workers then tell co-nationals about the firm in later years. The firm's visibility to that group, q_{kjt}, rises permanently in response to the shock. The shifter ξ_{kjt} then has a component that is more persistent than Assumption 4.4 allows, and it is positively correlated with the instrument. The direction of the resulting bias is unambiguous. Employment growth that is in fact driven by greater visibility is attributed to the wage and to the firm's share of the local market, so $\hat{\beta}_k$ and $\hat{\tilde{\sigma}}_{kg}$ are biased upward. The implied supply elasticity is then too large, and the implied markdown is too small. Because referral networks are stronger among immigrants than among natives (Munshi 2003; Dustmann et al. 2016), the bias is larger for immigrant types. If networks matter, the estimated immigrant–native gap in markdowns is therefore a lower bound on the true gap, and the share of the pay gap attributed to labor market power is conservative.

G Identification: moment conditions, proofs and details

G.1 Moment conditions for the labor supply parameters

The estimator of Equation (12) imposes

$$\mathbb{E}\Big[\Delta z_{kjt}\Big(\Delta_{e,e'}\log\tfrac{s_{kjt}}{s_{k0t}}-\beta_k\Delta_{e,e'}\log\tfrac{w_{kjt}}{w_{k0t}}-\tilde{\sigma}_{kg}\Delta_{e,e'}\log s_{kjt|g}\Big)\,\Big|\,S_{kjt}=1\Big]=0,\quad z\in\mathcal{Z},\tag{34}$$

for $e\geq 2$ and $e'\geq 3$, where $S_{kjt}=1$ denotes a firm–type cell with positive type-k employment in both $t-e'$ and $t+e$. These moments are equivalent to regressions of long differences in log employment shares on long differences in log wages and log inside shares, instrumented by short differences in the productivity proxies. With three instruments and two parameters per (k,g) cell the system is overidentified; in practice it is estimated by 2SLS with $e=2$ and $e'=3$.

G.2 Validity of the internal instruments

Throughout, $\Delta z_{kjt}\equiv z_{kjt}-z_{kjt-1}$ for $z_{kjt}\in\mathcal{Z}$ and $\Delta_{e,e'}x_t\equiv x_{t+e}-x_{t-e'}$.

Proof of Lemma 4.2. Under Assumption 4.4 the permanent term $\bar{\xi}_{kj}$ differences out, so $\Delta_{e,e'}\xi_{kjt}=(u_{kjt+e}+\psi_k u_{kjt+e-1})-(u_{kjt-e'}+\psi_k u_{kjt-e'-1})$, a function of u dated in $\mathcal{S}_\xi=\{t+e,\,t+e-1\}\cup\{t-e',\,t-e'-1\}$. Under Assumption 4.3, z_{kjt} is a function of period-t fundamentals and hence of u dated t and $t-1$, and z_{kjt-1} of u dated $t-1$ and $t-2$; so Δz_{kjt} is a function of u dated in $\mathcal{S}_z=\{t,\,t-1,\,t-2\}$. Setting $e\geq 2$ places the earliest lead date at $t+e-1\geq t+1$, strictly after t; setting $e'\geq 3$ places the latest lag date at $t-e'\leq t-3$, strictly before $t-2$. Hence $\mathcal{S}_\xi\cap\mathcal{S}_z=\varnothing$, and since u is serially uncorrelated the covariance vanishes. □

Three remarks. First, the two window lengths bind for different reasons: $e\geq 2$ keeps the lead endpoint clear of t, while $e'\geq 3$ keeps the lag endpoint clear of $t-2$, which is as far back as the MA(1) in Assumption 4.4 pushes the instrument. The window $e=2$, $e'=3$ is therefore the shortest admissible one, and is the window used by both LMS (Lamadon, Mogstad, and Setzler 2022) and CKMM (Chan et al. 2024). If $\psi_k=0$ the conclusion holds for $e\geq 1$ and $e'\geq 2$; if the departure from a permanent shifter is MA(q) it requires $e\geq q+1$ and $e'\geq q+2$. Second, no orthogonality between u and η is used anywhere: the covariance vanishes because the date sets are disjoint, so the conclusion holds even if the two are perfectly correlated at every date. Third, the proof places no restriction whatever on the behavior of equilibrium objects at t, $t-1$ and $t-2$, which is why movement in the raw series at those event times is consistent with the design.

Interpreting the two primitives. The permanent term $\bar{\xi}_{kj}$ carries the economics of Assumption 4.4: how attractive a given employer is to Economic Class immigrants from Asia, net of what

it pays and of how many it already employs, is a stable property of that firm–type pair, close to what Sorkin (2018) assumes. The disturbance u_{kjt} is specification and measurement error, chiefly the fact that the inside share $s_{kjt|g}$ the model treats as a population object is computed from finitely many workers, so the observed share is a noisy version of the predicted one and the discrepancy enters (9) exactly where ξ_{kjt} does (Berry 1994; Berry, Linton, and Pakes 2004). What is left unrestricted matters more. First, u_{kjt} may be arbitrarily correlated with η_{kjt} at every date, including contemporaneously: exogeneity comes from transitoriness, not independence. Second, $\bar{\xi}_{kj}$ is unrestricted, so persistent differences in which employers immigrants and natives find attractive are permitted without qualification — settlement patterns, the history of immigrant arrival in a place, workplaces differing in how readily a newcomer can navigate them, or differences in the treatment a group expects to receive.

Relevance. Assumption 4.2 delivers a first stage that survives the long difference but decays in the length of the lead window.

Proposition G.1. *Let Assumptions 4.2, 4.3 and 4.4 hold, and let $z_{kjt} \in \mathcal{Z}$. Then*

$$\mathbb{E}\big[\Delta z_{kjt} \cdot \Delta_{e,e'} \log \tilde{\gamma}_{kjt}\big] = \sum_{h=0}^{2} \delta_k^{e+h}\, \mathbb{E}\big[\Delta z_{kjt}\, \eta_{kjt-h}\big] - \big(1 - \delta_k^{e+e'}\big)\, \mathbb{E}\big[\Delta z_{kjt} \log \tilde{\gamma}_{kjt-e'}\big],$$

which is nonzero under Assumption 4.2 and declines geometrically in e at rate δ_k.

Proof. Iterating (10) from $t - e'$ to $t + e$ gives $\Delta_{e,e'} \log \tilde{\gamma}_{kjt} = -(1 - \delta_k^{e+e'}) \log \tilde{\gamma}_{kjt-e'} + \sum_{h=0}^{e+e'-1} \delta_k^h \eta_{kjt+e-h}$. By the proof of Lemma 4.2 the instrument is a function of innovations dated t, $t-1$ and $t-2$, which carry weights δ_k^e, δ_k^{e+1} and δ_k^{e+2} in that sum. □

Widening the lead window therefore buys robustness — it tolerates a longer-lived departure from a permanent shifter — at the cost of the first stage.

G.3 Threats to identification: testable implications

Assumptions 4.2 to 4.4 make three claims about timing: that the instrument is news at t and carries no information about innovations dated $t-3$ or earlier; that the productivity shock it captures is still present at the lead endpoint $t+2$; and that any correlation between productivity shocks and shocks to the labor supply shifter is transitory. Each has a counterpart in the data: a strong first stage, no pre-trends, a persistent productivity gap, and no continuing adjustment after the shock.

Relevance: the first stage. Assumption 4.2 requires $\sigma^2_{\eta_k} > 0$ and $\delta_k > 0$: type-specific productivity must actually move, and the movement must persist long enough for a short change

dated t to predict a long change spanning $t-3$ to $t+2$. This is directly testable, and it is the one identifying restriction that the data can reject outright. Table A8 reports partial F-statistics for each endogenous regressor and each of the 26 types. Most exceed 10 and the majority exceed 100, and every statistic for $\tilde{\sigma}_{kg}$ exceeds 50. Three types have an F-statistic below 10 for β_k—family-class and refugee women born in Africa, and refugee women born in the Americas—together accounting for less than 1 percent of the AKM sample and less than 3 percent of immigrants; the results are similar if these three are dropped or pooled with adjacent categories.[4] The event studies below provide the same evidence in a different form, by showing that a large share of the impact productivity difference survives to the lead endpoint.

Construction. First, I construct a discrete analog to a shock to firm-wide scale θ_{jt}: in each calendar year t I order firms by $\Delta \log R_{jt}$, split them into "treatment" (above median) and "control" (below median) groups, and follow the firm's overall earnings and employment for each. Second, I construct a discrete analog to a shock to type-specific productivity $\tilde{\gamma}_{kjt}$: for each worker type k, I repeat the exercise with revenue per type-k worker $\Delta \log(R_{jt}/l_{kjt})$ as the split variable and type-k earnings and employment as the outcomes. In both cases outcomes are plotted at event time $e = 0$ and in the years around it, normalized to zero at $e = -3$, the base period of the long difference, with the exercise performed separately by calendar year (and, in the second case, by type) and stacked. The figures then plot the event studies of the raw data, without controls or fixed effects of any kind. The vertical dashed lines at $e = -3$ and $e = +2$ mark the endpoints of $\Delta_{e,e'}$; the shaded band covers $e \in \{-2, \ldots, +1\}$, the dates the instrument is itself built from, at which the design places no restriction on the raw series.

What the figures can and cannot establish. The figures speak to the *timing* of shocks and of the adjustments that follow them, and in that capacity they provide evidence bearing on the identifying assumptions. They cannot, however, be mapped onto model parameters in the way their counterparts in LMS can, because firms here are not atomistic. A productivity innovation at a treated firm raises the wages its rivals post, and some of those rivals sit in the control group; the observed divergence therefore mixes the direct effect of the shock with its propagation through the market. Propagation is nonetheless a contemporaneous phenomenon: a rival responds to an innovation only once it has occurred, so no amount of strategic interaction can manufacture or conceal a differential trend in the years before the event, or generate one that persists after the shock itself has died out. The figures remain informative about timing, which is what the assumptions restrict.

4. Because Equation (12) has two endogenous regressors, joint significance is not sufficient: the instruments must move the wage and the inside share along non-proportional rays, as Section 4.5.2 explains. Table A8 therefore reports statistics for each endogenous regressor separately.

Firm-wide shocks. Figure A3 reports the split on $\Delta \log R_{jt}$. By construction the two groups differ in revenue growth from $t-1$ to t: treated firms experience about 18 log points more. Away from that year they are indistinguishable to the left of the band, with a treatment–control difference in log revenue of -0.004 at $e=-5$ and -0.002 at $e=-4$ and in log earnings of 0.001 at both dates. The gap that opens at impact does not close: log revenue is 0.160 apart at $e=+2$ and 0.154 at $e=+4$, 89 and 86 percent of the impact difference. This is the direct analogue of Figure 1 in LMS and validates the revenue instrument. The type-specific figures below carry the rest of the argument, since it is shocks to $\tilde{\gamma}_{kjt}$ that identify the type-specific parameters.

Type-specific shocks. Figures A4 and A5 report productivity and earnings for each type separately. Three patterns emerge, one for each claim.

No pre-trends. Averaging across types, the treatment–control difference in log revenue per type-k worker is 0.004 at $e=-5$ and 0.003 at $e=-4$, against an impact difference of 0.414; the corresponding differences in log type-k earnings are -0.002 at both dates against 0.051 at impact. This holds type by type and not merely on average: the $e=-4$ productivity difference lies in $[-0.016, 0.013]$ across types and the $e=-4$ earnings difference in $[-0.007, 0.002]$. Firms acting on innovations before they occur, or whose standing with type k were drifting rather than fixed, would separate here. This is not what the figures show.

A persistent productivity gap. The average productivity difference is 0.308 at $e=+2$, roughly three-quarters of the impact difference. The share surviving to the lead endpoint lies in $[0.53, 0.85]$ across types, with a median of 0.75, so relevance holds for every type and not only in aggregate. The instrument remains correlated with long changes in wages and shares at $t+2$, which is what Assumption 4.2 requires.

No catch-up. Beyond the lead endpoint the two groups neither converge nor continue to diverge. The average productivity difference moves from 0.308 at $e=+2$ to 0.292 at $e=+4$ and the earnings difference from 0.038 to 0.035, so the gap is stable after $t+2$. This is the pattern Assumption 4.4 predicts. A durable non-wage response to the productivity shock—a firm becoming permanently more attractive to type-k workers for reasons unmediated by the wage—would show up as continuing adjustment in the years after the shock, with type-k earnings drifting further apart while productivity held steady.

Sign of the residual bias. Should the MA(1) of Assumption 4.4 nonetheless fail—if referral networks make a firm's standing with type k respond durably to a productivity shock, so that ξ_{kjt} contains a component more persistent than MA(1) and $\mathrm{Cov}(\Delta_{e,e'}\xi_{kjt}, \Delta z_{kjt}) > 0$—the bias runs against the paper's finding. A positive covariance loads employment growth that is in truth a supply-shifter response onto the wage and inside-share terms of Equation (9), biasing $\hat{\beta}_k$ and $\hat{\tilde{\sigma}}_{kg}$ upward. By Equation (5) the firm-specific elasticity $\mathcal{E}_{kjt}$ is then overstated and, by Equation (8),

the markdown is understated. Referral networks of this kind are widely documented to be stronger among immigrants than among natives (Munshi 2003; Dustmann et al. 2016), so the bias should be larger for immigrant types. The estimated immigrant–native gap in markdowns is therefore a lower bound on the true gap, and the share of the pay gap attributed to labor market power is conservative.

G.4 Demand-side details

Quasi-differencing. Because ω_{khjt} inherits the AR(1) structure of Equation (10), it satisfies $\omega_{khjt} = \delta\,\omega_{khjt-1} + \zeta_{khjt}$. Subtracting δ times the lagged version of Equation (15) from the current version therefore removes ω_{khjt-1} entirely and leaves the innovation ζ_{khjt} as the only unobservable. This is a *quasi*-difference rather than a first difference because the lag is weighted by the estimated persistence: $\delta = 1$ recovers the first difference, $\delta = 0$ the level equation, and intermediate values partially difference the data. The construction is the standard one in dynamic panel estimation of production functions (Blundell and Bond 1998; Ackerberg, Caves, and Frazer 2015), and it is what makes variables dated $t-1$ valid instruments, since ζ_{khjt} is dated t and innovations are serially uncorrelated.

The intercept in Equation (16). Quasi-differencing (15) leaves an intercept $(1-\delta)\log(\rho_k/\rho_h)$, a constant given the pair. Relative curvature therefore enters only the intercept, while each ρ_k is identified from the coefficient on its own employment count. In estimation I allow this intercept to vary by year, $c_{khjt} \equiv c_{kht}$, absorbing both the pair-specific constant and any aggregate component of the relative productivity of types k and h in year t; no information in the intercept is used to identify $\{\rho_k\}$. Perfect substitutability is the joint null $\rho_k = 1$ for all k, which in the parameterization below is the linear restriction $\beta_k^1 = 0$ for all k, and it is testable.

Stacking across pairs. Equation (16) holds for each of the $K(K-1)/2$ pairs (k,h), each delivering an estimate of two curvature parameters and of δ, with no reason to prefer any particular subset; because not every firm employs every type, any chosen subset would arbitrarily discard the contribution of some firms, and estimating one pair at a time would also discard the cross-equation restriction that a given ρ_k is the same parameter wherever it appears. I therefore follow the multi-equation GMM approach of CKMM (Online Appendix D.4), stacking the moment conditions implied by Equation (16) across all available pairs, firms and years and estimating the system jointly.

Equation (16) is nonlinear in $(\{\rho_k\}, \delta)$. Rather than confront that nonlinearity directly, which CKMM note is close to intractable at this dimension, I follow them in treating the system as

linear in the compound parameters

$$\beta_k^1 \equiv \rho_k - 1, \qquad \beta_k^2 \equiv -\delta\,(\rho_k - 1), \qquad \delta, \tag{35}$$

so that the estimating equation becomes, for the pair (k, h),

$$y_{khjt} = \beta_k^1 \log l_{kjt} - \beta_h^1 \log l_{hjt} + \beta_k^2 \log l_{kjt-1} - \beta_h^2 \log l_{hjt-1} + \delta\, y_{khjt-1} + c_{kht} + \zeta_{khjt}, \tag{36}$$

which is linear in the $2K + 1$ parameters $\{\beta_k^1, \beta_k^2\}_{k \in \mathcal{K}}$ and δ. The system is estimated by linear GMM, equivalently two-stage least squares, and the structural parameters are recovered post-estimation as $\hat{\rho}_k = 1 + \hat{\beta}_k^1$. A pair–firm–year observation enters only if the firm employs both types in t and in $t - 1$, and the c_{kht} are absorbed by pair-by-year fixed effects.

The regressors dated $t - 1$ in (36) are predetermined with respect to ζ_{khjt} and instrument themselves; the contemporaneous employment counts are the endogenous regressors. The excluded instruments are the squares of the two lagged employment counts, each entered with the sign of the type it belongs to (K columns); the interaction of each lagged employment count with the lagged marginal revenue product ratio (K columns); the square of that lagged ratio; and the cross-product of the two lagged employment counts. The $K + 1$ predetermined regressors—the lagged ratio and the two lagged employment counts—instrument themselves. This gives $3K + 3$ instruments for the $2K + 1$ parameters, so the system is overidentified by $K + 2$ restrictions.

Specification test. The compound parameterization (35) is estimated without imposing the cross-equation restriction that links β_k^2 to the product of δ and β_k^1, which yields a specification test essentially for free. Each type's curvature can be recovered twice, from the contemporaneous coefficient as $1 + \hat{\beta}_k^1$ and from the lagged coefficient as $1 - \hat{\beta}_k^2 / \hat{\delta}$, and the two should agree if the AR(1) structure in Equation (10) and the common-δ restriction are correct. I report both and the discrepancy between them in Appendix I; the headline estimates and everything downstream use $1 + \hat{\beta}_k^1$.

H Estimation Appendix

H.1 Employment-weighted averages

Every parameter of the model other than β_k and ρ_k varies at the firm–type–year level, while the pay gap the paper decomposes is an average over workers. Averages reported in the main text and in the tables below are therefore employment-weighted throughout. For any firm–type–year object x_{kjt} and any set of cells $\mathcal{A}$—all cells, all native cells, all cells of a given admission category, all cells in a given province—the reported average is

$$\bar{x}_{\mathcal{A}} \;\equiv\; \frac{\sum_{(k,j,t)\in\mathcal{A}} l_{kjt}\, x_{kjt}}{\sum_{(k,j,t)\in\mathcal{A}} l_{kjt}}, \tag{37}$$

where l_{kjt} is type-k employment at firm j in period t in the AKM sample. The "average immigrant markdown" is thus the markdown faced by the average immigrant *worker*, not by the average firm that employs immigrants, and the same convention applies to elasticities, marginal revenue products, returns to scale and the concentration indices. Note that an average of a product is not the product of the averages, so quantities that decompose exactly within a firm–type–year cell need not do so after aggregation; however, discrepancies are small.

H.2 Construction of the Bootstrap Confidence Intervals

Notation and the resampling scheme. Let θ denote a scalar object of interest (an elasticity, a markdown, or an average of these) and let $\hat{\theta}$ be the estimate produced by applying the full estimation procedure of Section 4.6 to the observed sample. A bootstrap sample is drawn by resampling *employers*, stratified by labor market: within each market g, I draw, with replacement, as many firms as the market contains in the data, so that every resample preserves the number of employers per market. A firm that is drawn enters with its entire panel of observations. Market and within-nest employment shares s_{kjt} and $s_{kjt|g}$ are then recomputed on the resample rather than carried over from the original data, since it is precisely the sampling variation in these shares that the procedure is designed to capture. Repeating this B times and re-running the entire estimation procedure on each resample—including the first stage and the recovery of (β_k, σ_{kg})—yields the replications

$$\hat{\theta}^*_1,\ \hat{\theta}^*_2,\ \ldots,\ \hat{\theta}^*_B, \tag{38}$$

where $\hat{\theta}^*_b$ is the estimate of θ computed on the bth resample by exactly the steps that produced $\hat{\theta}$ on the observed sample. I write $\hat{\theta}^*$ for a generic element of Equation (38), that is, for the estimator viewed as a random variable over resamples, and q^*_p for the pth empirical quantile of $\{\hat{\theta}^*_b\}_{b=1}^{B}$. I use $B = 500$ in practice.

Step 1: the infeasible exact interval. Define the estimation error

$$R \;\equiv\; \hat{\theta} - \theta, \tag{39}$$

and let r_p denote the pth quantile of the distribution of R. By construction,

$$\Pr\Big(r_{\alpha/2} \;\le\; \hat{\theta} - \theta \;\le\; r_{1-\alpha/2}\Big) \;=\; 1 - \alpha. \tag{40}$$

Solving the event in Equation (40) for θ rather than for $\hat{\theta}$ gives

$$\Pr\Big(\hat{\theta} - r_{1-\alpha/2} \;\le\; \theta \;\le\; \hat{\theta} - r_{\alpha/2}\Big) \;=\; 1 - \alpha. \tag{41}$$

Equation (41) is an exact $1-\alpha$ interval for θ. Two features are worth noting. First, subtracting θ and $\hat{\theta}$ from Equation (40) reverses the direction of both inequalities, so the *upper* quantile of the error distribution determines the *lower* endpoint of the interval and vice versa. Second, the interval is not centered on $\hat{\theta}$ unless R is symmetrically distributed: if the estimator is biased upward, so that R is centered above zero, both endpoints shift down. The interval is infeasible because $r_{\alpha/2}$ and $r_{1-\alpha/2}$ depend on the unknown distribution of R.

Step 2: the bootstrap approximation. The bootstrap principle replaces the distribution of $\hat{\theta} - \theta$, which is unobservable because θ is unknown, with the distribution of its sample analogue

$$R^* \;\equiv\; \hat{\theta}^* - \hat{\theta}, \tag{42}$$

which is observable because it treats the estimate $\hat{\theta}$ as the population value and each replication $\hat{\theta}_b^*$ as a draw of the estimator around it. Writing q_p^* for the pth quantile of $\{\hat{\theta}_b^*\}_{b=1}^B$ and noting that $\hat{\theta}$ is a constant across replications,

$$r_p \;\approx\; p\text{th quantile of } R^* \;=\; q_p^* - \hat{\theta}. \tag{43}$$

Step 3: the reported interval. Substituting Equation (43) into Equation (41) yields the interval of Hall (1992),

$$\Big[\hat{\theta} - \big(q_{1-\alpha/2}^* - \hat{\theta}\big),\;\; \hat{\theta} - \big(q_{\alpha/2}^* - \hat{\theta}\big)\Big] \;=\; \Big[2\hat{\theta} - q_{1-\alpha/2}^*,\;\; 2\hat{\theta} - q_{\alpha/2}^*\Big], \tag{44}$$

which is what I report, with $\alpha = 0.05$, $B = 500$, and q_p^* the empirical pth quantile of the replications. In the notation of Equation (44) the interval is the reflection of the bootstrap distribution through the point estimate: each endpoint is $\hat{\theta}$ minus the distance from $\hat{\theta}$ to the opposite bootstrap quantile.

Relation to the percentile interval. The percentile interval $[q^*_{\alpha/2},\, q^*_{1-\alpha/2}]$ arises from a different approximation, in which the distribution of $\hat{\theta}^*$ stands in for the distribution of θ directly rather than for the distribution of the error. The results are similar using percentile intervals (available upon request).

Standard errors. I also report $\widehat{\text{se}}(\hat{\theta}) = \left[(B-1)^{-1}\sum_b(\hat{\theta}^*_b - \bar{\theta}^*)^2\right]^{1/2}$, where $\bar{\theta}^* = B^{-1}\sum_b \hat{\theta}^*_b$.

I Additional Results

I.1 Estimated labor supply parameters

Section 5.1 reports the firm-specific labor supply elasticity and the markdown it implies. This subsection reports the two primitives that generate them: β_k and σ_{kg}.

First stage diagnostics. Table A8 reports partial F-statistics for each endogenous regressor and each of the 26 types. Most exceed 10 and the majority exceed 100, and every statistic for $\tilde{\sigma}_{kg}$ exceeds 50. Three types have an F-statistic below 10 for β_k—family-class and refugee women born in Africa, and refugee women born in the Americas—together accounting for less than 1 percent of the AKM sample and less than 3 percent of immigrants; the results are similar if these three are dropped or pooled with adjacent categories.[5]

The two primitives move in opposite directions. Table A9 reports both by immigrant status and admission category. Immigrants have a *higher* β_k than natives, 0.697 (SE = 0.017) against 0.526 (SE = 0.006), and a substantially *lower* σ_{kg}, 6.81 (SE = 0.16) against 11.73 (SE = 0.18); the confidence intervals do not overlap in either case. The elasticity facing an employer small enough to ignore its own market share is the product $\beta_k\sigma_{kg}$, reported in column (3): 4.75 for immigrants against 6.17 for natives. The within-market substitution parameter dominates and the employment-margin parameter partly offsets it. The elasticity gap reported in Section 5.1 is thus a σ result, not a β result.

The ordering across admission categories follows the same pattern. The products $\beta_k\sigma_{kg}$ are 5.49 for economic-class immigrants, 4.06 for family-class immigrants, and 3.59 for refugees, and in each case the ranking is driven by σ_{kg} (7.57, 6.39, 4.97) rather than by β_k (0.725, 0.636, 0.723). The same holds for the gender comparison among natives: σ_{kg} is 13.94 for men against 7.63 for women, while β_k runs the other way, 0.427 against 0.708.

Interpretation. It is worth being precise about what this does and does not establish. A low σ_{kg} says that a wage cut at one employer sends a smaller share of that employer's workers to *other employers in the same market* than it would for a group with a high σ_{kg}. It is a statement about revealed substitutability, not about its source. Section 4.3 and Appendix F.2 set out several mechanisms consistent with it: thinner information about local job opportunities, language barriers in search, etc. These are observationally equivalent in the firm's problem, which depends on the slope of supply and not on why it has that slope. The counterfactuals of Section 6

5. Because Equation (12) has two endogenous regressors, joint significance is not sufficient: the instruments must move the wage and the inside share along non-proportional rays, as Section 4.5.2 explains. Table A8 therefore reports statistics for each endogenous regressor separately.

inherit this agnosticism: they ask what the pay gap would be if immigrants substituted across employers as natives do, without asserting what would have to change for that to happen.

A higher β_k for immigrants is worth noting in its own right, since it runs against a reading in which immigrants are simply less responsive to pay. Conditional on the set of employers they treat as alternatives, immigrants are *more* responsive to relative wages than natives are. What differs is the breadth of that set. Several of the type-level estimates of β_k exceed one, which is admissible: β_k is the elasticity of the employment margin with respect to pay and is not bounded above by unity.

Estimates by worker type. Table A10 reports β_k and Table A11 the employment-weighted average σ_{kg} for each of the 26 worker types, with the corresponding elasticities and markdowns in Tables A12 and A13. The type-level results show that the immigrant–native contrast is not driven by any single origin region or admission category: σ_{kg} is below the native value for every one of the 24 immigrant types.

I.2 Least-squares estimates of the supply parameters

Table A14 compares the instrumental-variable estimates of Section 5.1 with least-squares estimates of the same specification, and reports the comparison by worker type. The average elasticity falls from 5.25 to 0.289, and the implied share of the marginal revenue product withheld rises from 17.6 to 78.5 percent. For family-class immigrants the least-squares elasticity is negative, which lies outside the parameter space of Assumption 4.1, and for refugees it implies a markdown of essentially 100 percent. The direction of the bias is what the identification problem of Section 4.5 predicts: employers offering more attractive non-wage conditions attract more workers of a given type while needing to pay less to do so, so the least-squares slope is biased downward and the implied employer power correspondingly upward.

I.3 Estimated production technology parameters

Section 5.2 reports α_{jt} and ρ_k and explains why the levels of $\tilde{\theta}_{jt}$ and γ_{kjt} should not be read as a productivity ranking. This subsection reports all four and expands on that point.

The normalization. Equation (6) involves $\tilde{\theta}_{jt}$ and γ_{kjt} only through the product $\tilde{\theta}_{jt}\gamma_{kjt}$. They are separated by the restriction $\sum_{k \in C_{jt}} \gamma_{kjt} = 1$, imposed for every employer in every period. The restriction is a normalization, not a substantive assumption, but it has a consequence for cross-group comparisons of either parameter in levels.

Consider two employers producing identical output from identical inputs, one drawing on two worker types and one on eight. The adding-up restriction forces the second employer's γ_{kjt} to

average one-eighth against the first's one-half, and its $\tilde{\theta}_{jt}$ must be correspondingly larger to rationalize the same output. Nothing distinguishes the two technologies, yet the second appears to have higher total factor productivity and lower match-specific productivity on every type it employs.

Section 5.2 notes that this binds in these data, because immigrants work at larger employers and larger employers draw on more of the 26 types. The magnitude can be read off columns (3) and (4) of Table A15: the immigrant–native difference in $\log \tilde{\theta}_{jt}$ is 0.690 and the ratio of average γ_{kjt} is 0.41. I report the estimates for completeness but refrain from drawing any strong conclusions related to the immigrant-native pay gap from the estimated differences in these two parameters.

Estimates. Table A15 reports all four parameters. Average returns to scale are 0.291, with immigrant and native averages of 0.295 and 0.290; medians are lower than means in both cases (0.234 and 0.235), reflecting a right-skewed distribution across employers. The curvature parameter ρ_k averages 0.888 and is higher for immigrant labor (0.930) than for native labor (0.878). Tables A16, A17, A18, and A19 report the four parameters for each of the 26 worker types.

On ρ_k and the markdown. Restricting $\rho_k = 1$, as in Lamadon, Mogstad, and Setzler (2022) and much of the adjacent literature, makes the technology linear in each type's own input, so a type's marginal product at a given employer is independent of how many of that type the employer hires. Relaxing the restriction matters here because employers of immigrants employ immigrants intensively: if immigrant labor ran into diminishing returns more quickly, the marginal product of immigrants at those employers would be depressed for a purely technological reason, and part of what Section 5.1 attributes to wage-setting power would be misattributed. The estimates rule this out. The technology is if anything less curved in immigrant labor, so the correction runs against the paper's thesis rather than in favor of it, and is small.

The parameter is identified separately from the markdown. ρ_k enters the marginal revenue product and is recovered from the demand-side moment conditions of Appendix G.4, while the markdown is a function of the supply elasticity alone by Equation (8); the two enter the first-order condition multiplicatively and are estimated from separate sources of variation.

Specification test. The compound parameterization (35) is estimated without imposing the cross-equation restriction that links β_k^2 to the product of δ and β_k^1, which yields a specification test essentially for free. Each type's curvature can be recovered twice: from the contemporaneous coefficient as $\hat{\rho}_k = 1 + \hat{\beta}_k^1$, the baseline, and from the lagged coefficient as $\tilde{\rho}_k = 1 - \hat{\beta}_k^2/\hat{\delta}$, the alternative. If the AR(1) structure in Equation (10) and the restriction that δ is common across types are correct, the two should agree.

Table A17 reports both for all 26 types, which are similar across the two specifications. The median absolute discrepancy across types is 0.013 and the mean is 0.014, on estimates whose distance from perfect substitutability averages roughly five times that. The two sets rank the types almost identically, with a Spearman correlation of 0.982 and a Pearson correlation of 0.992. Most importantly for what follows, the substantive conclusion is unchanged: every alternative estimate also lies below unity, the largest being 0.981, so the joint null of perfect substitutability is rejected under either calculation of ρ_k, and the ordering of types by curvature that drives the counterfactuals in Section 6 is preserved.

I.4 Exposure to labor market concentration

I.4.1 The Generalized Concentration Index: More details

Section 5.3 defines the generalized concentration index of CKMM in Equation (17) and reports it by group. This appendix records the properties of the index that the main text uses without proving, relates it to the descriptive measures of Fact 4, and states the welfare result.

I.4.1.1 Bounds and limiting cases Write the two factors of Equation (17) as WGCI_{kt} and BGCI_{kt}. Each lies in $(0, 1]$, so $GCI_{kt} \in (0, 1]$ and a higher value denotes a more concentrated environment. Both attain one only in the degenerate case: for BGCI_{kt}, when every type-k worker is in a single market; for WGCI_{kt}, when within each market every type-k worker is at a single employer. Both approach zero as employment spreads across arbitrarily many markets, or across arbitrarily many employers within each. The indices are reported multiplied by 100 in Table 5 and Table A20.

Three features of the GCI are relevant:

1. As $\sigma_{kg} \to \infty$ in every market, the exponent s_{kgt}/σ_{kg} goes to zero and $\text{WGCI}_{kt} \to 1$: the index reduces to its between-market factor alone. Employers inside a market are perfect substitutes for the group, so how employment is distributed among them is irrelevant to the constraint the group faces, however few of them there are.

2. As $\sigma_{kg} \to 1$, the exponent is s_{kgt} and the within-market term enters at full weight. In this limit the index is a preference-weighted analogue of the employment-weighted average of within-market indices in Section 3, computed with entropies rather than sums of squares.

3. Holding σ_{kg} fixed, GCI_{kt} is increasing in a mean-preserving spread of the within-market employment shares $s_{kjt|g}$ and in a mean-preserving spread of the across-market shares s_{kgt}, so the index responds to concentration at both margins in the same direction. This is what makes the two-factor split in Equation (17) interpretable as a decomposition rather than an arbitrary product.

I.4.1.2 Welfare interpretation CKMM show that in the nested-logit class of Assumption 4.1 the expected utility of a type-k worker is decreasing in GCI_{kt}, holding the distribution of employer productivities fixed. The mechanism is the one in Equation (5): a more concentrated environment raises the average employment share of the firms a worker is likely to encounter, which lowers the elasticity each of them faces and raises the markdown each applies. Two qualifications attach to the statement. It holds the productivity distribution fixed, so it is a statement about the wage-setting consequences of market structure. And it is an ordinal comparison within a type: the index is not on a common utility scale across types, so the immigrant–native ratio of 1.87 reported in Section 5.3 says that immigrants are more exposed, not by how much their welfare is lower.

The HHI carries no such mapping. It is a summary of market structure that becomes a statement about worker welfare only under auxiliary assumptions about substitutability, and Section 5.3 shows those assumptions fail in exactly the direction that matters for the immigrant–native comparison: immigrants look less exposed than natives on within-market employer shares and more exposed once those shares are weighted by how substitutable the employers actually are.

I.4.1.3 GCI results Section 5.3 Table 5 reports the index and its two components by immigrant status and admission category, and Table A20 reports them for each of the 26 worker types.

The between-market component depends only on how a group's employment is distributed across local labor markets and involves no estimated parameter, so it carries no sampling uncertainty from the bootstrap; all of the sampling variation in the overall index enters through σ_{kg} in the within-market component. This is why the standard errors in Table 5 are an order of magnitude smaller on the between-market column.

Additionally, the reported within- and between-market averages do not multiply to the reported overall average. This is the aggregation issue of Equation (37): each column is an employment-weighted average of a type-year index, and the average of a product is not the product of the averages. Within each worker type and year the decomposition is exact. The discrepancy is 3.4 percent for immigrants and under 1 percent for natives.

I.5 Correlates of the labor supply shifters

This appendix section asks the following question: which observable employer characteristics move with the supply shifter ξ_{kjt}, and are there differences between immigrants and natives?

Recovering the shifter. Given $\hat{\beta}_k$ and $\hat{\sigma}_{kg}$, Equation (4) inverts to

$$\widehat{\log \xi_{kjt}} = \log \frac{s_{kjt}}{s_{k0t}} - \hat{\beta}_k \log \frac{w_{kjt}}{w_{k0t}} - (1 - 1/\hat{\sigma}_{kg}) \log s_{kjt|g}, \tag{45}$$

which recovers, for each type–employer–year cell, the average non-wage payoff that rationalizes observed employment given observed pay and market share. I then estimate

$$\widehat{\log \xi_{kjt}} = X'_{jt}\beta^u + \psi^u_n + \psi^u_p + e^u_{kjt}, \tag{46}$$

where X_{jt} collects observable employer characteristics (revenue, headcount, total wage bill), ψ^u_n are two-digit NAICS industry effects, and ψ^u_p are province effects. Equation (46) is estimated separately for immigrants and natives; Figure A6 reports the estimated β^u and fixed effects by immigrant status.

$\widehat{\log \xi_{kjt}}$ is a residual: the part of a worker type's attachment to an employer not accounted for by pay or by the employer's size in the market. It is consistent with amenities in the usual sense, but equally with information, credential recognition, referral networks, or anything else that makes a job more likely to be taken at a given wage. The observationally equivalent information-based model of Appendix F.2 generates the same object with no reference to preferences. See Section 5.4 for more details.

Incremental R^2. The estimated fixed effects are normalized against an omitted category, so their levels are not comparable across specifications. To compare explanatory power instead, I group the right-hand side into three blocks—location effects, industry effects, and other employer characteristics—and compute, for each, the drop in R^2 from removing it from the full specification, $\Delta R^2 \equiv R^2_{(1)} - R^2_{(2)}$.

Figure A7 reports the three increments, with results aggregated by immigrant status.

Location effects account for a larger share of the variation in $\widehat{\log \xi_{kjt}}$ for immigrants than for natives: 0.219 against 0.126, averaging across types. This is the same phenomenon as the between-market concentration of Section 5.3, viewed from the supply side—immigrants' attachment to employers is more strongly organized by where those employers are—and it is the largest single difference between the two groups. It is also ordered across admission categories in the way the settlement literature would predict, at 0.234 for the economic class and 0.248 for the family class but only 0.173 for refugees, whose initial placement is determined by resettlement policy rather than by an existing community. Industry effects run the other way and by less, 0.094 for immigrants against 0.139 for natives, and the observable employer characteristics are close to indistinguishable, 0.097 against 0.107. The full specification explains more of the variation for immigrants than for natives, $R^2 = 0.489$ against 0.389, and essentially all of the difference is

the location block. The patterns are consistent with the literature on enclave settlement (Card 2001; Altonji and Card 1991) and on occupational mismatch among immigrants (Aydede and Dar 2016). They are correlations with a recovered residual and nothing in the model's identification rests on them.

Workplace risk. As a further descriptive check, I estimate Equation (46) separately for each type and regress the resulting industry effects on the rate of workplace illness or injury in that industry:

$$\hat{\psi^u_{kn}} = \eta_0 + \eta_1 x_n + \nu^u_{kn}. \tag{47}$$

Column (1) of Table A21 shows that industries with higher rates of illness or injury have lower estimated $\log \xi_{kjt}$: workers are less attached to riskier industries at a given wage, which is what one would expect. Column (2) tests whether this differs between immigrants and natives and finds no significant difference. Immigrants are more likely to hold dangerous jobs (Lay et al. 2018), but the estimates give no indication that this reflects differences in supply. Data on illness and injury rates are from the U.S. Bureau of Labor Statistics and are matched to Canadian industries at the two-digit NAICS level.

I.6 Heterogeneity across provinces

There is considerable heterogeneity in labor supply elasticities and markdowns across provinces, with lower elasticities and higher markdowns in Quebec, the Prairies, and Atlantic Canada (Table A22). This is intuitive: workers in Ontario and British Columbia have more suitable job alternatives than workers elsewhere.

Two patterns are worth noting. First, the immigrant–native ordering holds in every province group—immigrants face a lower elasticity and a larger markdown everywhere—so the aggregate result of Section 5.1 is not driven by any single region. Second, the size of the gap varies considerably. In the Prairies and Atlantic provinces, immigrants face markdowns 15 percentage points higher on average than the native-born; in Quebec the gap is 10 percentage points; in Ontario and British Columbia it falls to 4 and 5 percentage points respectively.

J Counterfactual Appendix

J.1 Solving for the counterfactual equilibrium

The fixed point is solved by under-relaxed Jacobi iteration. Given a wage vector $w^{(n)}$, each firm's first-order condition is evaluated at the implied employment allocation to obtain a target wage $\tilde{w}^{(n)}$, and the next iterate is the convex combination

$$w^{(n+1)} \;=\; \lambda\, \tilde{w}^{(n)} \;+\; (1-\lambda)\, w^{(n)}, \qquad \lambda = 0.1. \tag{48}$$

The under-relaxation weight λ is set well below one because each firm's best response depends on the wages posted by its competitors, and full updating produces oscillation rather than convergence. Iteration stops when $\|w^{(n+1)} - w^{(n)}\|_\infty < \$500$ in annual earnings, so the criterion binds on the single worst-adjusting worker-firm cell rather than on an average. Every counterfactual reported in Section 6 satisfies this criterion in fewer than 500 iterations.

Because the criterion is a sup norm, the aggregate object of interest is far better behaved than the stopping rule requires. The employment-weighted pay gap of Equation (18) is stable over the iterations preceding termination in every counterfactual: once the largest remaining wage adjustment anywhere in the economy is below \$500, the gap typically ceases to move at the precision reported in the tables. The results are therefore not sensitive to the exact point at which the algorithm is stopped.

J.2 The counterfactual sample and solver validation

Sample. The counterfactuals are computed on 2016 alone, for the reason given in Section 6.2: the fixed point must be solved once per counterfactual, and fifteen solutions on the full panel are not feasible.

Validation. At the estimated parameters, with no restriction imposed, the solver should return the observed equilibrium regardless of where it is started. I therefore set the wage for every firm and type to the sample average —a point far from the data—and iterate to convergence. The resulting immigrant–native pay gap is 15.27 percent, against an observed gap in 2016 of 14.99 percent. The remaining 0.28 percentage points reflect the stopping rule of Equation (48). Every counterfactual in Section 6 is solved by the same algorithm from the same starting point.

J.3 Marginal contributions across elimination orders

Table A23 summarizes the Shapley decomposition discussed in Section 6, and Table A24 reports the underlying 32 contributions in full, that is, for each block b, the change in the gap from

eliminating b given each of the eight subsets $\mathcal{C} \subseteq \mathcal{B} \setminus \{b\}$ already eliminated. Read down a panel, the entries show how much one policy lever accomplishes depending on which other levers are already in place.

Employer substitutability is the largest contributor under every one of the eight conditioning sets, but its contribution is roughly twice as large when skill mix and technology are left at their estimated values (-13.10 to -16.98 pp) as when they have been equalized (-8.31 to -8.54 pp). Strategic interactions run the other way: eliminating them closes little on its own, or once $\mathcal{R}$ or $\mathcal{V}$ has been removed (-0.66, -0.05 and -0.61 pp), but a great deal (between 4.29 and 4.43 pp) once employers have been made equally substitutable and skill mix and technology are left intact. Both patterns follow from Equation (5), in which shares and $\beta_k \sigma_{kg}$ enter multiplicatively. In policy terms, the two levers that act on wage-setting power are complements: improving information and mobility raises what antitrust enforcement can accomplish, because it moves immigrants toward firms that hold larger shares of the markets they hire from.

Skill mix and technology is the only block whose contribution changes sign depending on the other blocks. Equalizing it closes the gap by 4.94 to 5.80 pp when employers remain differentially substitutable across groups, but *widens* it, by up to 3.74 pp, once they do not. The mechanism is the reallocation described in Section 6.3. When wage-setting power is intact, immigrants are held at low productivity employers, and equalizing productive capacity raises what they earn there. Once wage-setting power has been removed, immigrants have already sorted toward more productive firms, and equalizing technology across firms flattens the productivity ladder they climbed, returning part of that gain. The two levers are therefore substitutes: selection and industrial policy deliver most where wage-setting power is unaddressed, and least where it has been addressed already. This is the counterpart to the complementarity in the preceding paragraph, and neither pattern is visible in a decomposition that holds the allocation of workers across firms fixed.

Appendix References

Abowd, JM, F Kramarz, and DN Margolis. 1999. "High wage workers and high wage firms." *Econometrica* 67 (2): 251–333.

Ackerberg, Daniel A., Kevin Caves, and Garth Frazer. 2015. "Identification Properties of Recent Production Function Estimators." *Econometrica* 83 (6): 2411–2451.

Altonji, Joseph G., and David Card. 1991. "The Effects of Immigration on the Labor Market Outcomes of Less-skilled Natives." In *Immigration, Trade, and the Labor Market.* University of Chicago Press.

Aydede, Yigit, and Ather Dar. 2016. "The Cost of Immigrants' Occupational Mismatch and the Effectiveness of Postarrival Policies in Canada." *IZA Journal of Migration* 5 (1): 9.

Berry, Steven, Oliver B. Linton, and Ariel Pakes. 2004. "Limit Theorems for Estimating the Parameters of Differentiated Product Demand Systems." *The Review of Economic Studies* 71 (3): 613–654.

Berry, Steven T. 1994. "Estimating Discrete-Choice Models of Product Differentiation." *The RAND Journal of Economics* 25 (2): 242–262.

Blundell, Richard, and Stephen Bond. 1998. "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models." *Journal of Econometrics* 87 (1): 115–143.

Card, David. 2001. "Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration." *Journal of Labor Economics* 19 (1): 22–64.

Card, David, Ana Rute Cardoso, Joerg Heining, and Patrick Kline. 2018. "Firms and Labor Market Inequality: Evidence and Some Theory." *Journal of Labor Economics* 36 (1): S13–S70.

Card, David, Ana Rute Cardoso, and Patrick 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.

Chan, Mons, Kory Kroft, Elena Mattana, and Ismael Mourifie. 2024. "An Empirical Framework For Maching with Imperfect Competition."

Citizenship and Immigration Canada. 2011. "Evaluation of the Provincial Nominee Program." *Evaluation Report Ci4-75/2011E-PDF,* accessed November 5, 2025.

Dostie, Benoit, Jiang Li, David Card, and Daniel Parent. 2023. "Employer policies and the immigrantânative earnings gap." *Journal of Econometrics* 233 (2): 544–567.

Dustmann, Christian, Albrecht Glitz, Uta Schönberg, and Herbert Brücker. 2016. "Referral-based Job Search Networks." *The Review of Economic Studies* 83 (2): 514–546.

Fosgerau, Mogens, Emerson Melo, André de Palma, and Matthew Shum. 2020. "Discrete Choice and Rational Inattention: A General Equivalence Result." *International Economic Review* 61 (4): 1569–1589.

Hall, Peter. 1992. *The Bootstrap and Edgeworth Expansion.* Springer Series in Statistics. New York, NY: Springer.

Kroft, Kory, Isaac Norwich, Matthew Notowidgdo, and Stephen Tino. 2025. "The Labor Market Returns to Permanent Residency." *NBER Working Paper No. 34630.*

Lamadon, Thibaut, Magne Mogstad, and Bradley Setzler. 2022. "Imperfect Competition, Compensating Differentials, and Rent Sharing in the US Labor Market." *American Economic Review* 112 (1): 169–212.

Lay, A. Morgan, Agnieszka Kosny, Anjana Aery, Karl Flecker, and Peter M. Smith. 2018. "The occupational health and safety vulnerability of recent immigrants accessing settlement services." *Canadian Journal of Public Health* 109 (3): 303–311.

Matějka, Filip, and Alisdair McKay. 2015. "Rational Inattention to Discrete Choices: A New Foundation for the Multinomial Logit Model." *American Economic Review* 105 (1): 272–298.

McFadden, Daniel. 1978. "Modeling the Choice of Residential Location." In *Spatial Interaction Theory and Residential Location,* edited by A Karlqvist, L Lundqvist, J Weibull, and F Snickars, 75–96. Amsterdam: North Holland.

Munshi, Kaivan. 2003. "Networks in the Modern Economy: Mexican Migrants in the U.S. Labor Market." *The Quarterly Journal of Economics* 118 (2): 549–599.

Sorkin, Isaac. 2018. "Ranking Firms Using Revealed Preference." *Quarterly Journal of Economics* 133 (3): 1331–1393.