How Restricting Migrants' Job Options Affects Both Migrants and Existing Residents*

Wilbur Townsend† and Corey Allan‡

September 23, 2025

Do restrictive work visas protect existing residents' wages? We study New Zealand's 'Essential Skills' visa, which required migrants work only for firms that could not recruit New Zealanders. Loosening a single migrant's restrictions does not affect their wages. However, loosening the restrictions on *all* an occupation's migrants does increase wages. These results are consistent with a wage-posting model where each firm pays migrants and residents equally. We estimate such a model. The restrictions decreased migrants' average wage by 5.6%. Though most residents were unaffected, 2.5% had their wage decreased by more than 2%.

*For advice and comments we thank Sydnee Caldwell, Raj Chetty, Sahil Chinoy, Shani Cohen, Prasanthi Cottingham, Zoe Cullen, Oren Danieli, Christian Dustmann, Larry Katz, Pat Kline, Kory Kroft, Sarah Gao, Jeff Gortmaker, Quan Le, Shakked Noy, Ben Schoefer, Jesse Shapiro, Jesse Silbert, Clare Suter, Shreya Tandon, Chris Walters, and seminar participants at Motu, Harvard, Chicago, Toronto, Singapore Management University, NYU, Berkeley, Michigan and Brookings. For help obtaining our data we thank Nafees Anwar, Andrew Hoy, David Paterson, Jan-Yves Ruzicka, Jess Williamson and other staff from MBIE and Stats NZ.

These results are not official statistics. They have been created for research purposes from the Integrated Data Infrastructure (IDI) which is carefully managed by Stats NZ. For more information about the IDI please visit `https://www.stats.govt.nz/integrated-data/`. The results are based in part on tax data supplied by Inland Revenue to Stats NZ under the Tax Administration Act 1994 for statistical purposes. Any discussion of data limitations or weaknesses is in the context of using the IDI for statistical purposes, and is not related to the data's ability to support Inland Revenue's core operational requirements.

This research was supported by the Institute for Humane Studies (grant IHS017729), by the Michael S. Chae Macroeconomic Policy Fund, and by the James N. Bailey Fellowship Fund.

†University of California, Berkeley. Contact: wilbur.townsend@gmail.com.

‡Ministry of Business, Innovation and Employment.

1 Introduction

Governments typically restrict the jobs in which international migrants can work. For example, a migrant's visa conditions might require that they work in a 'high-skill' occupation, that they work in a certain region, or that they work for a certain firm.[1] These restrictions aim to provide firms with migrant labor while protecting existing residents' wages.[2] However, by limiting migrants' outside job options, these restrictions may allow employers to pay migrants less. Moreover, if employers pay migrants and residents equally, these restrictions may reduce residents' wages as well.

This paper asks how restricting migrants' job options affects the migrants themselves, the existing residents of their new country, and employers. We study New Zealand's Essential Skills visa, which was New Zealand's primary employer-sponsored visa between 2008 and 2022. Workers holding an Essential Skills visa could only switch to a new employer if they obtained a new visa. Obtaining a new visa could be difficult. In most occupations, the visa was limited to employers who had demonstrated that they could not recruit New Zealanders. We evaluate the Essential Skills visa by using quasi-experimental analyses to discipline a structural model of the labor market.

In specifying our structural model, a critical modeling choice is how wages are set. If employers individually bargain with each worker, restricting migrants' job options could reduce migrants' wages, but residents would only be affected insofar as their productivity is affected. If instead employers 'post' wages that pay migrants and residents equally, the wages they choose will depend on the job options of prospective employees. Restricting migrants' job options would thus reduce not only migrants' own wages but also the wages of their resident colleagues. In sum, to accurately model how restricting migrants' job options affects wages we must accu-

[1]Examples of visas with a 'high-skill' restriction include the US H-1B visa, which is limited to occupations which require a "theoretical and practical application of a body of highly specialized knowledge", and the UK Skilled Worker visa, for which the British government maintains a list of occupations in which a migrant must work. In a similar vein, governments often require that prospective employers pass a 'labor market test' demonstrating that no existing residents are available to fill the job. An example of a regional restriction is that imposed by the Australian regional provisional visas, which require their holders live and work outside Australia's largest cities.

[2]The official intent of the US H1-B program is to "help employers who cannot otherwise obtain needed business skills and abilities from the U.S. workforce by authorizing the temporary employment of qualified individuals" while establishing "certain standards in order to protect similarly employed U.S. workers from being adversely affected by the employment of the nonimmigrant workers" `https://web.archive.org/web/20240829160808/https://www.dol.gov/agencies/whd/immigration/h1b`. The official objectives of New Zealand's Essential Skills visa included "helping New Zealand firms maintain capacity... while not displacing New Zealanders from employment opportunities or hindering improvements to wages" `https://web.archive.org/web/20230218070502/https://www.immigration.govt.nz/opsmanual/66920.htm`.

rately model how wages are set.

We use quasi-experimental analyses to understand how the firms that we study set wages. We study both individual-level shocks to specific migrants' job restrictions and market-level shocks that affect all migrants in an occupation. Loosening an individual migrant's job restrictions does not affect their earnings: when a migrant transitions to an unrestricted resident visa, they often switch jobs but, on average, receive no gain in earnings. In contrast, loosening the restrictions on all migrants in an occupation increases both job-switching and average earnings. These results are consistent with a wage-posting model in which each firm pays an equal wage to all of the workers in each occupation.

We thus evaluate the Essential Skills job restrictions using a wage-posting model. We compare equilibrium wages under the job restrictions to a counterfactual simulation in which migrants' job options are unrestricted. The restrictions decreased migrants' average wage by 5.6%, and decreased the wages of many residents as well. The restrictions increased aggregate profits, but these benefits were concentrated among a minority of firms: the median firm was actually hurt. The restrictions decreased aggregate welfare by $411m New Zealand dollars — equal to 42% of migrants' baseline earnings.

Our paper begins by formally showing that different wage-setting models make different predictions about the effects of restricting migrants' job options. We consider a framework in which firms have wage-setting power because workers have varying preferences over firms. Migrants can work at only some firms whereas residents can work at any. Within this framework, we consider three different wage-setting models, which each correspond to a different equilibrium concept. In a competitive equilibrium, firms choose employment levels taking wages as given. Each worker will be paid their marginal product, so job restrictions can only affect wages by reallocating workers to more or less productive firms. In a bargaining equilibrium, a worker's wage splits the difference between their outside job option and their marginal product. As such, restricting a migrant's job options can decrease their wages, even if their marginal product is unaffected. In a wage-posting equilibrium, each firm commits to a single wage that remunerates all its workers equally. Wage-posting firms account for the distribution of prospective workers' outside job options when choosing wages. As such, restricting migrants' job options can decrease all workers' wages.

To assess these models we first present descriptive wage regressions using a cross-section of full-time workers around the 2018 Population Census. On average, migrant workers who hold an Essential Skills visa earn about 20% less than residents. Most of this difference remains when we control for age, education and weekly hours of work. However, the difference disappears when we control for a worker's job: among workers who share both a firm and an occupation, those on an Essential Skills visa earn no less than others. In other words, in jobs with many migrants, both migrants and residents are paid unusually little.

We next ask how migrants' earnings change when they transition from an Essential Skills visa (which restricts their job options) to a resident visa (which does not). This *individual-level* shock allows us to assess the importance of individual-level bargaining: In a bargaining equilibrium, but not in our other wage-setting models, an individual worker's wage depends on her own outside job option. When a migrant transitions from an Essential Skills visa to a resident visa, her job options expand: many new residents change firms shortly after receiving their resident visa. However, we find that receiving a resident visa has no mean effect on a worker's earnings. Using both a matched-control design and a lottery design we rule out any effects on log earnings larger than 0.01 with 95% confidence. We continue to find no effect when we focus on managerial and professional occupations in which existing research has found bargaining to be more common (Caldwell & Harmon, 2019; Lachowska, Mas, Saggio, & Woodbury, 2022; Caldwell, Haegele, & Heining, 2024).

We also study how migrants' average hours of work change when they obtain a resident visa: we find no evidence that they are affected. We do find that new migrants disproportionately move into larger firms, suggesting that they typically use their newfound freedom to move to a higher-amenity job, rather than a higher-wage job.

In a bargaining equilibrium, expanding the job options of a worker will increase the wage of that worker *provided she stays at her original firm*. Of course, many new residents decide to switch to one of their new job options. Among these job-switchers, the effect of the expanded job options is ambiguous: while the new residents will have a stronger bargaining position, they may decide to move to a low-wage, high-amenity job. If they do so, the results discussed thus far could be consistent with a bargaining equilibrium.

To assess this 'job-switching' concern, we estimate the average effect of greater job options *among those workers who do not switch firms*. For this exercise we use only our sample of workers who entered a random resident visa lottery. The probability that a worker switches firms is itself affected by the lottery, and thus simply conditioning on workers who do not switch firms would yield a biased estimate of this conditional average effect. Nonetheless, under plausible assumptions, this 'stayers average treatment effect' can be set- or point-identified.

The intuition is as follows. We observe the realized earnings of lottery-winning job stayers. The identification challenge is to identify counterfactual earnings for lottery-winning job-stayers: what they would have earned, on average, had they lost the lottery. To do so, we must infer how counterfactual earnings for lottery-winning job-stayers differ from realized earnings among lottery losers. We consider three different methods for inferring this 'selection bias'. The first method trims the empirical distribution of earnings among lottery losers (Lee, 2009). The second method infers the sign of the selection bias using a shape restriction on the job-switching decision. The third method point-identifies the selection bias by estimating a parametric model of the job-switching decision.

Across all three methods we find no evidence that winning a resident visa lottery increases the wages of those who remain at their initial firm. This result is consistent with both competitive equilibrium and wage-posting equilibrium. It is inconsistent with individual bargaining playing a significant role in wage-setting.

Our analyses of individual-level shocks are consistent with both competitive equilibrium and wage-posting equilibrium. To distinguish *these* models, we next analyse a market-level shock. Wage-posting equilibrium predicts that a market-level loosening of migrants' job restrictions can increase migrants' wages — and indeed it can increase residents' wages as well. We test this prediction by studying how earnings are affected by an occupation's inclusion in an Essential Skills in Demand list: in occupations on these lists, migrants can work for any firm willing to employ them. Wage-posting equilibrium predicts that listing an occupation will increase earnings in that occupation. We confirm that prediction using a difference-in-difference research design: earnings in newly-listed occupations increase more quickly than do earnings in other occupations. These increased earnings coincide with increased worker mobility. We find no evidence that earnings effects can be explained by the reallocation of workers to more productive firms, as would need to be the case in a competitive equilibrium. Our analysis of market-level shocks is thus most consistent with wages being set in a wage-posting equilibrium.

Having corroborated a wage-posting model, we can use it to evaluate the distributional and welfare effects of migrant job restrictions. The model is identified under the assumption that shocks to the amenity value of employment are orthogonal to historic and contemporaneous productivity shocks. Our structural estimates imply that an atomistic firm will face a labor supply elasticity of about 0.8. Workers are largely unwilling to move between geographic locations: the labor supply elasticity to an atomistic location is about 0.1. Foreign-born workers have slightly less elastic labor supply than New Zealand-born workers. Firms have near-constant returns to scale and can easily substitute between workers in different occupations.

With these estimates we can calculate the equilibrium effect of the Essential Skills visa system. We compare equilibrium wages given the observed set of firms which can employ migrants to a counterfactual equilibrium in which all firms can employ migrants. The Essential Skills visa system reduces the average wage of an Essential Skills migrant by about 5.6%. These effects are heterogeneous: migrants who largely work with other migrants typically have their wage reduced by over 20%. The typical resident is unaffected — migrants make up a small proportion of the overall workforce — but about 2.6% of residents have their wage reduced by more than 2%. These residents' average wage is about 12% less than the average wage in the full population.

The Essential Skills visa system also increases aggregate firm profit — although it slightly decreases the profit of most firms, it substantially increases the profit of a few firms. Firms which benefit from the restrictions are on average larger, and typically employ migrants. There are, however, some firms which benefit from the restrictions despite employing no migrants — these firms benefit purely because the restrictions weaken competition between firms.

The overall annual welfare loss from the system is $411m NZD, which is equal to 42% of migrants' earnings. This welfare loss is due in part to a reduction in aggregate output, but is due mostly to the reduction in the non-pecuniary amenities which migrants receive from their firms. Recall our reduced-form result that migrants who transition to a resident visa frequently switch firms, but receive no gas abd in in earnings. By revealed preference, this result suggests that the non-pecuniary aspects of employment are important. Our structural model can quantify non-pecuniary amenities and indeed we find that they account for most of the cost of the Essential Skills visa system.

We conclude our paper by studying how the Essential Skills job restrictions relate to other aspects of immigration policy. Specifically, we use our structural estimates to ask both how additional migration would affect residents' wages, and whether migration would affect residents more were migrants' job options unrestricted. We find that restricting migrants' job options accentuates the negative wage effects of migration, sometimes greatly. A migration shock which increased the total number of bus drivers by 10% would decrease the mean wage of resident bus drivers by 2.7%. Were migrant bus drivers' job options unrestricted, such a shock would decrease the mean wage of resident bus drivers by only 1%. The Essential Skills visa increases firms' power over migrants, encouraging lower wages for both migrant bus drivers and their resident colleagues.

Literature review. We contribute to three literatures. First, we contribute to the literature evaluating visa policy.[3] The existing literature on work visas has used reduced-form techniques to either study how market-level shocks to a visa system affect migrants' wages (Naidu, Nyarko, & Wang, 2016; Ahrens, Beerli, Hangartner, Kurer, & Siegenthaler, 2023), or to ask how an individual worker's wages change when she transitions away from a restrictive visa (Wang, 2021).[4] Analyses of market-level shocks have found much larger effects than has the analysis of individual-level transitions. By studying both market-level shocks and individual-level shocks, we are able to reconcile these results: in a wage-posting equilibrium, an individual worker's visa status might matter little even though the overall visa system matters a lot. We further extend

[3]This literature is distinct from the larger literature asking how migration affects the labor market (Dustmann, Schönberg, & Stuhler, 2017; Borjas, 2003; Card, 1990).

[4]A closely related literature studies the effects of regularizing undocumented migrants on the migrants themselves and on existing residents (Kossoudji & Cobb-Clark, 2002; Borjas & Edo, 2023; Elias, Monras, & Vázquez-Grenno, 2025; Di Porto, Martino, & Naticchioni, 2018; Carrozzo, 2022).

this literature by studying the effects of migrant restrictions both on residents and on firms.[5]

Second, we contribute to the literature assessing models of wage-setting.[6] The classic assumption is that wages are set in competitive equilibrium: firms take wages as given and choose employment levels accordingly. Among models that depict firms choosing wages, some depict firms bargaining with each of their workers individually (Stole & Zwiebel, 1996; Pissarides, 2000; Bagger, Fontaine, Postel-Vinay, & Robin, 2014; Caldwell & Danieli, 2024) while others depict firms posting inflexible wage policies (Burdett & Mortensen, 1998; Manning, 2005; Card, Cardoso, Heining, & Kline, 2018). The distinction between these three models of wage-setting is critical for understanding the causes of wage inequality and for assessing whether firms' labor market power results in a misallocation of workers (Berger, Herkenhoff, & Mongey, 2022; Silbert & Townsend, 2024). Different wage setting models also imply different effects of migration on residents' earnings (Amior & Manning, 2021; Amior & Stuhler, 2024; Chassamboulli & Palivos, 2013, 2014; Chassamboulli & Peri, 2015; Battisti, Felbermayr, Peri, & Poutvaara, 2018; Albert, 2021; Borjas, 1995, 2013; Monras, 2020).

In surveys, about a third of workers and firms report bargaining over their wage (Brenzel, Gartner, & Schnabel, 2014; Hall & Krueger, 2012).[7] However it remains an open question how important such bargaining is in practice. This paper joins a growing literature that assesses the empirical importance of wage bargaining by studying the effects of outside job options (Caldwell & Danieli, 2024; Caldwell & Harmon, 2019; Ahrens et al., 2023; Lachowska et al., 2022). We extend the literature on outside options in two ways. First, we study both individual-level and market-level shocks to workers' outside options to separately test wage bargaining and wage posting. Second, we distinguish bargaining effects from job-switching effects by estimating the effect of better outside options among those workers who do not switch firms.[8]

Third, we contribute to the literature estimating structural models of labor market power.

[5] Amior and Manning (2021) and Amior and Stuhler (2024) argue that migrants may have less elastic labor supply, in part because their job options are restricted. As such, increasing migration may harm existing residents by inducing firms to set less generous wages. We ask how restricting migrants' job options mediates the effects of immigration in subsection 8.3. Doran, Gelber, and Isen (2022) ask how firms are affected by work visa policy by studying firm-level lotteries for the US H1-B visa.

[6] Beyond the outside options literature cited below, this literature includes analyses of rent-sharing (e.g. Carvalho, Galindo da Fonseca, and Santarrosa, 2023; Bloesch, Larsen, and Taska, 2022), analyses comparing workers' productivity to their wages (e.g. Delabastita and Rubens, 2022), and an analysis that requires productivity — estimated by inverting a labor supply system — be orthogonal to a concentration instrument (Roussille and Scuderi, 2024).

[7] Caldwell et al. (2024) report that firms often pay similarly productive workers different wages, even when explicit back-and-forth bargaining does not occur.

[8] This estimator exploits the panel structure of wage data, and thus is in a similar vein to those presented by *e.g.* Card and Hyslop (2005) and Taber and Vejlin (2020).

This literature consists of analyses which infer market power by estimating firm-specific labor supply elasticities (Manning, 2005; Lamadon, Mogstad, & Setzler, 2022; Azar, Berry, & Marinescu, 2022; Berger et al., 2022; Kroft, Luo, Mogstad, & Setzler, 2020; Roussille & Scuderi, 2023; Chan, Kroft, Mattana, & Mourifié, 2024), and an overlapping literature which compares wages to marginal products from an estimated production function (Yeh, Macaluso, & Hershbein, 2022; Delabastita & Rubens, 2022; Lamadon et al., 2022; Kroft et al., 2020).

We extend the structural labor literature in two ways. First, our model can explain how market concentration for one worker type (*e.g.* migrants) can affect wages for other workers (*e.g.* residents). To do so, it abandons the typical assumption that firms are 'strategically small'; in our model, a firm which employs most of the migrants in a labor market will realize that it faces a weaker elasticity of labor supply and so will pay a lower wage.

We also extend the structural labor literature by estimating production functions in which distinct occupations are distinct inputs. Existing analyses either treat labor as homogenous, or as varying in some low-dimensional way (e.g. 'high-skill' vs. 'low-skill'). We show how to parsimoniously estimate a CES production function in which over 800 occupations can enter separately. By estimating firms' ability to substitute across occupations we can infer whether residents in one occupation (*e.g.* building laborers) are affected by job restrictions in another occupation (*e.g.* carpenters). The key insight is that we can estimate the production function using, as the dependent variable, the marginal product of a given occupation in a given firm, rather than a firm-level financial variable like revenue or value-added. We infer these marginal products by inverting firms' wage-setting equations, given estimates of the firm-occupation-specific labor supply elasticities.[9]

Roadmap. Section 2 provides our theoretical framework. Section 3 presents some additional details on the Essential Skills visa system and describes our data. Section 4 describes the cross-sectional relationship between a worker's visa status, the visa status of her colleagues, and her earnings. Section 5 describes how workers' earnings change when they transition from an Essential Skills visa to a resident visa. Section 6 describes how earnings in an occupation are affected by the inclusion of that occupation on an Essential Skills in Demand list. In Section 7 we estimate our structural model. Section 8 uses the structural model to describe the equilibrium effects of the Essential Skills visa system and analyse immigration policy. Section 9 concludes.

[9]We follow Chan et al. (2024), as well as the empirical IO literature on differentiated products (Berry, Levinsohn, & Pakes, 1995), by estimating production functions using the firm's first order condition and by allowing for strategic interactions between firms. We extend the approach taken by Chan et al. (2024) by showing how strategic considerations generate spillovers across worker types, by allowing distinct occupations to play distinct roles in firms' production functions, and by identifying our model without using equilibrium objects, like revenue or market shares, as instruments.

2 Three Models of Wage-Setting

In this section we present three models of wage-setting. The models are identical except for their solution concept: competitive equilibrium, bargaining equilibrium, or wage-posting equilibrium. We show that these solution concepts disagree as to how restricting migrants' job options affects the labor market.

For simplicity, we assume in this section that workers have homogeneous productivity; they differ only in their preferences. In Appendix Section A we extend our model: we endow workers with varying (and, potentially, firm-specific) productivities, and we assign workers an occupation, which may be substitutes or complements and of which firms may make differing use. Our results continue to hold in this more general setting. Given that the model in Appendix A generalizes the model presented here, we include only proofs for the results in Appendix A.

2.1 The theoretical framework

We consider a static, partial-equilibrium, perfect-information labor market comprising a unit continuum of workers $\mathbf{I}$ and a finite set of firms $\mathbf{F}$. Each worker $i \in \mathbf{I}$ will be employed by a single firm $f_i \in \mathbf{F}$.

Production. Firms' only input is the total mass of labor they employ $L_f \equiv \int_{i:f_i=f} di$, with which they produce output using a concave and differentiable production function $y_f : \mathbb{R}^+ \to \mathbb{R}^+$. We focus only on strategic interactions in the labor market, and so assume that output is sold at unit price in a competitive market.

Labor supply. Each worker is either a *migrant* or a *resident*. Residents can be employed by any firm while migrants can only be employed by a subset of firms. Let $\mathbf{F}^{\text{migrant}} \subset \mathbf{F}$ denote the subset of firms at which migrants can work. Let $\mathbf{I}^{\text{migrant}} \subset \mathbf{I}$ denote the subset of workers who are migrants and let $\mathbf{I}^{\text{resident}} = \mathbf{I} \setminus \mathbf{I}^{\text{migrant}}$ denote the subset who are residents. For each worker i, let $\mathbf{F}_i \in \left\{\mathbf{F}^{\text{migrant}}, \mathbf{F}\right\}$ denote the set of firms at which she can work.

Workers have preferences over their wage and over the firm which employs them. We assume that the preferences of each worker $i \in \mathbf{I}$ over their wage w and firm f can be represented as

$$u_i(w, f) = we^{v_{i,f}},$$

where $v_{i,f} \in \mathbb{R}$ measures i's taste for working at f. We assume that the conditional distribution of the random vector $v_i \equiv \left(v_{i,f}\right)_{f \in \mathbf{F}}$, given a migrant's visa status, is absolutely continuous, with a differentiable PDF.

For each firm f, let $L_f\left(w_f; w_{-f}\right)$ be the mass of workers who would choose firm f, were it to pay all its workers wage w_f and other firms were to pay wages $w_{-f} \equiv \left(w_g\right)_{g \in \mathbf{F} \setminus \{f\}}$:

$$L_f\left(w_f; w_{-f}\right) \equiv \mathbb{P}\left[f \in \underset{g \in \mathbf{F}_i}{\arg\max}\left\{u_i\left(w_g, g\right)\right\}\right].$$

(By the absolute continuity of the distribution of v_i, measure zero workers will be indifferent between firms.) We assume that each firm f's labor supply function $L_f\left(w_f; w_{-f}\right)$ is concave in its own wage.

Defining equilibrium. We compare three solution concepts. In each solution concept we require that wages be strictly positive.

A *competitive equilibrium* comprises an assignment of workers to firms $\left(f_i\right)_{i \in \mathbf{I}}$ and a firm-specific wage schedule $\left(w_f\right)_{f \in \mathbf{F}}$ such that each worker's firm yields her the maximal utility from her choice set:

$$\forall i \in \mathbf{I}: f_i \in \underset{f \in \mathbf{F}_i}{\arg\max}\, u_i\left(w_f, f\right),$$

and the mass of labor assigned to each firm maximizes its profits, taking wages as fixed:

$$\forall f \in \mathbf{F}: L_f\left(w_f; w_{-f}\right) \in \underset{L \in \mathbf{R}^+}{\arg\max}\left\{y_f\left(L\right) - w_f L\right\}.$$

A *bargaining equilibrium* comprises an assignment of workers to firms $\left(f_i\right)_{i \in \mathbf{I}}$ and a worker-specific wage schedule $\left(w_i\right)_{i \in \mathbf{I}}$ such that each worker is assigned to that firm which could yield her maximal utility, were she paid her marginal product

$$\forall i \in \mathbf{I}: f_i \in \underset{f \in \mathbf{F}_i}{\arg\max}\, u_i\left(y_f'\left(L_f\right), f\right),$$

and her wage is bargained *a la* Nash (1950):

$$\forall i \in \mathbf{I}: w_i \in \underset{w \in \mathbf{R}^+}{\arg\max}\left\{\left(u_i\left(w, f_i\right) - u_i^{\text{option}}\right)^{\beta}\left(y_{f_i}'\left(L_{f_i}\right) - w\right)^{1-\beta}\right\}, \tag{1}$$

where $\beta \in (0,1)$ measures workers' bargaining power and u_i^{option} is the maximal utility that i could receive at an alternative firm:

$$\forall i \in \mathbf{I}: u_i^{\text{option}} \equiv \max_{f \in \mathbf{F}_i \setminus \{f_i\}} u_i\left(y_f'\left(L_f\right), f\right).$$

In particular, we follow existing bargaining models by assuming that a worker would counterfactually receive her marginal product, were she employed at her outside option. Cahuc, Postel-Vinay, and Robin (2006) provide an extensive-form justification for this approach using a Rubinstein (1982) alternating-offers bargaining game.

A *wage-posting equilibrium* comprises an assignment of workers to firms $\left(f_i\right)_{i\in\mathbf{I}}$ and a firm-specific wage schedule $\left(w_f\right)_{f\in\mathbf{F}}$ such that each worker's firm yields her the maximal utility from her choice set:

$$\forall i \in \mathbf{I} : f_i \in \underset{f\in\mathbf{F}_i}{\arg\max}\, u_i\left(w_f, f\right),$$

and each firm's wage is chosen to maximize its profits, given the wages set by the other firms:

$$\forall f \in \mathbf{F} : w_f \in \underset{w\in\mathbf{R}^+}{\arg\max}\left\{y_f\left(L_f\left(w; w_{-f}\right)\right) - wL_f\left(w; w_{-f}\right)\right\}.$$

We assume that there exists a unique competitive equilibrium wage schedule, a unique bargaining equilibrium wage schedule, and a unique wage-posting equilibrium wage schedule. For restrictions on workers' preferences that yield a unique wage-posting equilibrium, see *e.g.* Caplin and Nalebuff (1991); Chan et al. (2024).

2.2 Predictions for wage-setting

We will show that our candidate wage-setting models have differing implications for how shocks to workers' outside options affect workers' wages. Before doing so, we characterize equilibrium wages.

Competitive wages equal the marginal product of labor:

Lemma 1. *Under a competitive equilibrium, each firm $f \in \mathbf{F}$ will pay a wage equal to its marginal product: $w_f = y'_f\left(L_f\right)$.*

Bargained wages will depend both on the worker's actual marginal product of labor and on the marginal product of labor at the worker's outside option:

Lemma 2. *Under a bargaining equilibrium, each worker $i \in \mathbf{I}$ will receive wage*

$$w_i = \beta y'_{f_i} + (1-\beta)\exp\left(v_{i,f_i^{option}} - v_{i,f_i}\right) y'_{f_i^{option}},$$

where f_i^{option} is the worker's outside option:

$$f_i^{option} \in \underset{f\in\mathbf{F}_i\setminus\{f_i\}}{\arg\max}\, u_i\left(y'_f, f\right).$$

A firm's posted wages will be marked down from its marginal product by a factor that depends on its firm-specific elasticity of labor supply:

Lemma 3. *Under a wage-posting equilibrium, each firm $f \in \mathbf{F}$ will pay wage*

$$w_f = \left(\frac{1}{1+\frac{1}{\eta_f}}\right) y_f'$$

where η_f is the firm-specific elasticity of labor supply:

$$\eta_f = \frac{w_f}{L_f} \frac{\partial L_f \left(w_f; w_{-f}\right)}{\partial w_f}.$$

With these characterizations, we can study shocks to workers' outside options. We first consider a comparative static with respect to the job options of an individual worker: we consider turning a migrant into a resident. In all of our solution concepts, such a shock could either decrease or increase the worker's earnings if she decided to move firms. However, our solution concepts disagree as to how such a shock would affect a worker who decided to remain at her original firm. Neither her firm's marginal product nor its labor supply elasticity would be affected by the job options of an atomistic worker, and thus (by lemmas 1 and 3) under both a competitive equilibrium and a wage-posting equilibrium, the wage received by a worker who did not switch firms would be unchanged. In contrast, under a bargaining equilibrium, a worker's wage depends on her own outside option; when her outside option strengthens her wage will increase:

Proposition 1. *Consider a migrant i whose bargaining equilibrium firm f_i would be unchanged, were she a resident:*

$$f_i = \underset{f \in \mathbf{F}^{migrant}}{\arg\max}\, u_i \left(y_f'\left(L_f\right), f\right) \in \underset{f \in \mathbf{F}}{\arg\max}\, u_i \left(y_f'\left(L_f\right), f\right).$$

Removing her from the set of migrants $\mathbf{I}^{migrant}$ would weakly increase her bargaining equilibrium wage. Moreover, her bargaining equilibrium wage would strictly increase provided that her new outside option is better than her old, i.e. provided that

$$\max_{f \in \mathbf{F} \setminus \{f_i\}} u_i \left(y_f'\left(L_f\right), f\right) > \max_{f \in \mathbf{F}^{migrant} \setminus \{f_i\}} u_i \left(y_f'\left(L_f\right), f\right).$$

We also consider a market-level shock: allowing migrants to work at any firm. We will show that such a shock will have an unambiguous effect on posted wages under the following three assumptions.

First, we assume that firms' marginal products are homogeneous and constant:

Assumption 1. *Firms have a homogeneous production function which features constant returns to scale:*

$$\exists y : \mathbb{R}^+ \to \mathbb{R}^+ \text{ such that } \forall f \in \mathbf{F} : y_f = y \text{ and } \forall L_a, L_b \in \mathbb{R}^+ : y'\left(L_a\right) = y'\left(L_b\right).$$

Assumption 1 shuts down two potential channels through which a market-level shock might affect wages: It assumes that firms have a homogeneous production function, foreclosing workers moving to more- or less-productive firms. Is also assumes that this production function exhibits constant returns to scale, foreclosing changes to each firm's marginal product. Of course, firms do not actually share a homogeneous, constant returns to scale production function. When in Section 6 we study a market-level shock to migrants' job options, we will assess whether our estimated effects are due to an increase in workers' marginal product.

Second, we assume that workers' preferences over firms are independent of workers' visa status:

Assumption 2. *Migrants' preferences and residents' preferences over firms share the same distribution:*

$$\forall x : \mathbb{P}\left[v_i \le x \,\middle|\, i \in \mathbf{I}^{migrant}\right] = \mathbb{P}\left[v_i \le x \,\middle|\, i \in \mathbf{I}^{resident}\right].$$

When Assumption 2 fails, migrants might have such strong preferences for some firm that allowing migrants to work at that firm would induce the firm to reduce its wage, even though doing so would mean that the firm lost its resident workers. When we specify our structural model in Section 7 we weaken Assumption 3 by requiring only that migrants and residents have the same distribution of preferences given an observable demographic type.

Given a wage schedule $\left(w_g\right)_{g\in\mathbf{F}}$, let $\eta_f^{\text{resident}}\left(\left(w_g\right)_{g\in\mathbf{F}}\right)$ denote the firm-specific elasticity of residents' labor supply:

$$\eta_f^{\text{resident}}\left(\left(w_g\right)_{g\in\mathbf{F}}\right) \equiv \frac{\partial}{\partial \log w_f} \log \mathbb{P}\left[f \in \operatorname*{argmax}_{f\in\mathbf{F}} u_i\left(w_f, f\right) \,\middle|\, i \in \mathbf{I}^{\text{resident}}\right].$$

(Recall our assumption that workers' preferences across firms $v_i \equiv \left(v_{i,f}\right)_{f\in\mathbf{F}}$ follow an absolutely continuous distribution, with a differentiable PDF. This assumption guarantees that the elasticities η_f^{resident} exist and are themselves differentiable.)

Assumption 3. *Each firm's elasticity of resident labor supply is increasing in other firms' wages:*

$$\forall \left(w_g\right)_{g\in\mathbf{F}}, \forall h \neq f \in \mathbf{F} : \frac{\partial}{\partial w_h} \eta_f^{resident}\left(\left(w_g\right)_{g\in\mathbf{F}}\right) > 0.$$

Assumption 3 is a supermodularity assumption. By Lemma 3, Assumption 3 implies that wages in wage-posting equilibrium are strategic complements: if one firm increases its wage, other firms will increase theirs in response.[10] Assumption 3 holds, for example, when preferences take the logit form $v_{i,f} \sim \text{Gumbel}\left(\delta_f, 1\right)$. Assumption 3 need not hold more generally. When preferences are bimodal, for example, one firm might respond to a rival's wage increase by 'ceding the

[10] We prove this claim within the proof of Proposition A.2.

ground', reducing its own wage and retaining only its most committed workers. The structural parametric form we impose in Section 7 does not imply Assumption 3.

With these assumptions in hand, we consider a comparative static with respect to the set of firms at which migrants can work.

Proposition 2. *Consider expanding* $\mathbf{F}^{migrant}$ *such that migrants can work at any firm:* $\mathbf{F}^{migrant} = \mathbf{F}$. *Under a competitive equilibrium, Assumption 1 implies that workers' wages will be unchanged. Under a wage-posting equilibrium, assumptions 1, 2 and 3 imply that wages at every firm will increase.*

Assumption 1 implies that firms' marginal product is exogenous. Under a competitive equilibrium, firms always pay their marginal product, and thus competitive equilibrium wages are unaffected by the shock.

In a wage-posting equilibrium, a firm will choose to pay a high wage if paying a lower wage would lose the firm many workers. When migrants' job choices are restricted, fewer of a firm's workers will be marginal, inducing the firm to set lower wages.

Proposition 2 shows that expanding migrants' outside options can increase wages for both migrants and residents. This result contrasts with the effects of expanding the job options of a single worker, which would have no effect on posted wages. In a wage-posting equilibrium, migrants' collective outside options matter in a way that their individual outside options do not.

Proposition 2 does not predict how expanding $\mathbf{F}^{\text{migrant}}$ would affect bargaining equilibrium wages. Surprisingly, this effect is ambiguous, even under Proposition 2's assumption that firms have a homogeneous and exogenous marginal product. Expanding a migrant's job options could allow them to work at a firm for which they have a strong (non-pecuniary) preference. Such a firm may be able to bargain a lower wage than could the migrant's initial firm. The model thus provides a theoretical rationalization for governments' claims that restricting migrants' job options can prevent migrant exploitation, further motivating the empirical analysis below.[11]

3 Institutional Context, Data and Summary Statistics

We study the period August 2008 to March 2022. During this period, the Essential Skills Work Visa was New Zealand's primary employer-sponsored work visa, constituting 20–40% of all temporary work visas held (working holiday visas, family visas and study-to-work visas made up the

[11]For example, the objectives of New Zealand's Accredited Employer work visa — the successor to the Essential Skills visa that we study — includes "reducing risks around business models and practices that might enable migrant exploitation" `https://web.archive.org/web/20230217034245/https://www.immigration.govt.nz/opsmanual/77094.htm`.

bulk of the remainder). Between 1% and 3% of all full-time workers held an Essential Skills visa over this period. Though Essential Skills migrants disproportionately worked in certain occupations, these occupations did not necessarily require a university education — as we discuss in more detail below, Essential Skills migrants often worked in the construction trades, in hospitality, and in agriculture.

An Essential Skills Work Visa tied its holder to a particular employer, and new visas were limited to jobs for which there were 'no New Zealanders available'. In most occupations, a prospective employer would have to convince an immigration official that they had made a "genuine attempt to attract and recruit" New Zealanders, for example by "advertising the vacancy in a national newspaper and/or website". This process was not merely perfunctory: about 10% of Essential Skills applications were declined. In contrast, in occupations included on an 'Essential Skills in Demand' list, immigration officials were instructed to assume that no New Zealanders were available. As such, migrants in these occupations could obtain a visa for any firm willing to employ them.

Unlike workers who held an Essential Skills visa, workers who held a resident visa could generally work for any firm willing to employ them. (Some resident visas required that holders remain in skilled employment for the first 3–12 months of the visa.) Resident visa holders could also stay in New Zealand indefinitely, were eligible for government welfare (*e.g.* unemployment benefits and subsidized education), and could vote. In this paper we refer to resident visa holders and citizens collectively as 'residents'.

Essential Skills visas were granted for between 1 and 5 years, depending on when they were granted and on the skill level of the migrant's occupation. Expiring visas could be renewed. Among workers who arrived in New Zealand during 2010 on an Essential Skills visa, 37% had transitioned to a resident visa by 2015; 8% remained on a temporary work visa while 55% had left New Zealand (New Zealand Productivity Commission, 2022). Migrants who transitioned to a resident visa typically obtained a Skilled Migrant Category visa, eligibility for which was based on a lengthy (and time-varying) formula. New Zealand work experience was one component of this formula so migrants often became eligible for a resident visa after spending some time in New Zealand.

During the period we study, about 18% of New Zealand employees were union members (OECD, 2021). When setting wages for non-unionised employees, firms were constrained only by a nation-wide minimum wage.[12] Over our sample period New Zealand's minimum wage grew from 59% of the median wage in 2008 to 70% of the median wage in 2022 (OECD, 2024). In 2020, 9% of employees received the minimum wage (Maré & Hyslop, 2021).

[12]Minimum wages were lower for some young workers and apprentices. A sectoral minimum wage schedule for care and support workers was introduced towards the end of our sample period.

3.1 Data

We use tax, immigration and survey data, linked within Statistics New Zealand's *Integrated Data Infrastructure*. Our analytic sample comprises all people aged at least 18; we restrict the sample further as appropriate (*e.g.* to those with positive earnings, or for whom we can infer an occupation).

Earnings and employers. New Zealand's income tax system requires that employers submit monthly wage records. We use this 'Employer Monthly Schedule' data to observe each worker's earnings and employers in each month. This data generally excludes owner-operators and the self-employed.

Visa status. We observe visa spells using administrative immigration records. One limitation is that we only observe visas; we don't observe citizenship. We assume that any worker in New Zealand who lacks a visa is a citizen. This is a plausible assumption given that there are no known illegal entries into New Zealand and that very few temporary entrants overstay their visa.[13] We also use immigration records to observe entries into resident visa lotteries.

Occupation. Essential Skills migrants were required to list their occupation in their visa application; we observe these occupations directly. We also observe administrative occupation records for many other visa-holders, including holders of some resident visas. Finally, we observe the occupation of almost all workers at the 2013 and 2018 population censuses. We impute occupations at the worker-month level using any occupation record within the prior 5 years: we prioritise records lexicographically, prioritising any immigration record over any census record, and then taking the most recent record prior to the month of interest. All occupation data is coded using the ANZSCO 6-digit classification.

Demographics and location. We use both administrative and census records to infer workers' age, gender, country of birth and education level. We measure workers' geographic location of residence using Statistics New Zealand's internal predictions, which draw on both administrative and census records.

Hours of work. We measure hours of work using the population censuses and the Household Labour Force Survey.

Confidentiality. Statistics New Zealand requires that we randomly round all counts to base 3.[14] Means are calculated using randomly-rounded denominators, and proportions are calculated using both randomly-rounded numerators and randomly-rounded denominators. All

[13]In 2016, 0.035% of temporary entrants overstayed their visa `https://web.archive.org/web/20231209062108/https://www.immigrationtrust.co.nz/new-zealand-immigration-news/overstayer-numbers-drop-to-new-low`.

[14]*I.e.* if $N \bmod 3 = 0$, N is left as is; if $N \bmod 3 = 1$, N is rounded up with probability $\frac{1}{3}$ and rounded down with probability $\frac{2}{3}$; if $N \bmod 3 = 2$, N is rounded up with probability $\frac{2}{3}$ and rounded down with probability $\frac{1}{3}$.

other values are left unadjusted.

3.2 Summary statistics

Our sample is described in Table 1. Essential Skills migrants comprise 1.3% of the workforce during the 2018 Census and comprise 1.4% of all worker-month observations in our sample. About 70% of Essential Skills migrants are men, and they are slightly more likely to live in Auckland than are other workers. At the time of the 2018 Population Census, the most common country of origin was the the Philippines, accounting for only about 27% of migrants; India, the UK, China, South Africa and Fiji are other common origin countries. Essential Skills migrants have varied educational backgrounds, with about 39% having only a high school education, 18% having a trade qualification, 24% having a bachelor's degree and 14% having a postgraduate degree. Essential Skills migrants are very likely to be paid employees (rather than *e.g.* self-employed or unpaid workers in a family business), are very likely to work full-time, and rarely hold a secondary job. Their earnings are somewhat lower than residents and citizens.[15]

Table 2 lists the 20 most common occupations for Essential Skills migrants during March 2018. These occupations are mostly semi-skilled, although the list includes both a highly skilled occupation (Resident Medical Officer, *i.e.* a junior doctor) and low-skilled occupations (*e.g.* sales assistant). In most of these occupations, Essential Skills migrants comprise a small-but-significant proportion of all workers. However, if we weight firm-by-occupation observations by the number of Essential Skills migrants therein — *i.e.*, we study the visa status of Essential Skills migrants' colleagues — we find that Essential Skills migrants very often work with other Essential Skills migrants. This segregation will play an important role in inducing firms which employ Essential Skills migrants to pay less generous wages.

4 The Cross-Sectional Relationship between a Worker's Visa Status, her Colleagues' Visa Status, and her Earnings

In this section we estimate cross-sectional Mincer-style regressions which examine the relationship between a worker's visa status and her earnings. Essential Skills migrants earn about 20% less than other full-time workers. However, this difference is entirely explained by the jobs in which Essential Skills migrants work: conditional on an occupation-by-firm fixed effect, Essential Skills migrants on average the same as other workers. Similarly, we show that both Essential

[15]The sample in March 2013, when the 2013 Population Census was held, looks similar — although between 2013 and 2018 the total number of Essential Skills migrants grew from 12894 to 21603. One difference is that in 2013 only 17% of Essential Skills migrants were born in the Philippines.

Skills migrants and other workers are paid less when more of their colleagues hold an Essential Skills visa.

Analytic sample. In this section we study the population of full-time employees with a consistent employer prior to either the 2013 or 2018 population census. Studying cross-sections around the censuses allows us to restrict our sample to full-time employees, and ensures that we observe the occupation of all workers. Specifically, our analytic sample comprises workers who, in the corresponding census, reported that they were working for a single employer for at least 30 hours per week. We include only workers who reported that they were paid employees, rather than *e.g.* self-employed. We additionally restrict our sample to workers who, in the linked tax data, have non-zero earnings and a unique employer in the 12 months prior to the census. We discard workers with imputed census data. Finally, we include only workers who hold an Essential Skills visa, a resident visa, or who are citizens; we exclude workers who hold other temporary visas.

Cross-sectional wage regressions. The first column of Table 3 reports a simple regression in which a worker's log earnings are regressed on an indicator for the worker holding an Essential Skills visa. Across both censuses, Essential Skills migrants earn about 20% less than other full-time workers.

Column (2) of Table 3 adds controls to the regression: log hours of work, indicators for education levels, and indicators for decadal age bins. (Coefficient values for these controls are reported in Appendix Table A1.) Adding these controls attenuates the coefficient on Essential Skills status by at most 0.03.

Latter columns of Table 3 add fixed effects to the regression equation: occupation, firm, or their interaction. Doing so substantially attenuates the coefficient on Essential Skills status. In the regression reported in Column (5), which includes both controls and an occupation-by-firm fixed effect, the coefficient on Essential Skills status is a precisely estimated zero.

Panels B and D of Appendix Table A1 repeat these regressions with a larger sample: workers with a unique employer in the 6 months prior to the census, rather than only those with a unique employer in the 12 months prior. Doing so expands the sample to include those with less stable employment, though it may also increase measurement error if, for example, workers receive large annual bonuses. Estimates using this sample are similar.

The relationship between a worker's earnings and her colleagues' visa status. The results reported in Table 3 suggest that Essential Skills migrants earn much less than other workers, but only because of the jobs they hold. In other words, both migrants and non-migrants will earn less when more of their colleagues in the same firm-by-occupation cell are migrants. Figure 1 tests this directly. We first calculate the proportion of workers in each firm-by-occupation cell who hold an Essential Skills visa. We then group workers, both by the proportion of Essential

Skills migrants in their cell (rounded to the nearest 10%), and by their own visa status.

Figure 1 depicts mean earnings for each of these groups. Both Essential Skills migrants and other workers earn less in cells containing more Essential Skills migrants. Comparing two workers on the same visa status – one who works almost exclusively with Essential Skills migrants and one who works almost exclusively with non-migrants, the former would tend to earn about $15000 less than the latter.

5 How Transitioning From a Migrant Visa to a Resident Visa affects a Worker's Earnings

In this section we follow workers who transition from an Essential Skills visa to a resident visa. Whereas an Essential Skills visa was restricted to a particular employer, a resident visa allowed its holder to work for any willing employer. We thus interpret a transition to a resident visa as an individual-level loosening of job restrictions.

In Subsection 5.1 we track the earnings of a large set of workers who transition from an Essential Skills visa to a resident visa. We compare these workers to those who worked in the same occupation and firm in the month before the migrant received a resident visa. New residents often move firms, suggesting that their job restrictions had been binding. We find no evidence that a resident visa increases their earnings.

The key threat to Subsection 5.1's matched-control identification strategy is individual-level shocks: If workers apply for resident visas in anticipation of negative shocks to their earnings which leave their colleagues' earnings unaffected (*e.g.* a health shock), the matched-control estimator won't recover the causal effect of a resident visa. To assess this threat, in Subsection 5.2 we study Essential Skills migrants who enter a random lottery for a resident visa. Though this sample is somewhat unrepresentative of the general population of Essential Skills migrants, we can use it to estimate the causal effect of a resident visa under weaker conditions. Comparing lottery-winners to lottery-losers confirms that a resident visa substantially affects job-switching but does not affect earnings.

In light of Proposition 1, these results seem to be inconsistent with bargaining models of wage-setting: workers' job options are expanding but their earnings are unaffected. However, Proposition 1 only speaks to wage changes *among workers who remain at their original firm*. Concretely, some new residents might be switching from high-wage low-amenity jobs to low-wage high-amenity jobs; our causal estimand could in principle average over a negative job-switching effect and a positive bargaining effect.

In Subsection 5.3 we isolate bargaining effects by estimating the average effect of winning a resident visa lottery *among new residents who stay at their baseline firm*. Among such job-stayers, winning the lottery need not be independent of potential outcomes. Nonetheless, we show that this 'stayers average treatment effect' can be estimated by exploiting the panel structure of our data. Under weak conditions, the stayers average treatment effect is set-identified; under stronger conditions this estimand is point-identified. Using various sets of assumptions we find no evidence that winning a resident visa lottery increases the average earnings of job-stayers. By Proposition 1, this result is inconsistent with the bargaining model of wage-setting.

5.1 Matched-control estimates

Analytic sample. In this subsection, we study workers who transitioned from an Essential Skills visa to a resident visa. We limit our sample to workers who held an Essential Skills visa for at least 18 months prior to receiving a resident visa. We include only those who can be assigned both an occupation and a unique firm in the month prior to them receiving a resident visa. Moreover, we include only workers who share this baseline firm-by-occupation cell with at least one other worker; we will use these *baseline colleagues* as a matched control for the new resident.[16]

Identifying variation. Typically, workers who transitioned from an Essential Skills visa to a resident visa obtained a 'Skilled Migrant Category Resident Visa'. Eligibility for this visa was based on a lengthy, time-varying formula. Inputs into this formula included the worker's qualifications, their partner's qualifications, their partner's English language ability, their New Zealand family support, and their work experience.

Migrants often became eligible for a resident visa after spending some time working in New Zealand, both because they accumulated additional points and because the required number of points varied over time. Given that eligibility changed over time, often with little warning, a migrant who had become eligible for a resident visa may not remain eligible for long. As such, migrants often applied for a resident visa as soon as they were eligible. Nonetheless, migrants did have some discretion about when they applied, and thus a natural concern is that they sometimes applied for a resident visa in anticipation of a shock to their earnings. In this subsection, our estimates are unaffected by any shocks that affect the new resident and their baseline colleagues equally. In the following subsection we will study only workers who obtain their resident visa through a random lottery; those estimates are additionally robust to shocks that affect the new resident and not their baseline colleagues.

Results. The blue series in Figure 2 Panel A depicts the proportion of new residents working

[16] We do not restrict the visa status of our matched control. As shown in Panel B of Appendix Figure A1, 12 months prior to the modal migrant receiving a resident visa 62% of the controls held a resident visa or were citizens. Four years later, 86% were. There is no discontinuity in the control group's visa status around the time when the modal worker becomes a resident.

for a firm for the first time, by the number of months before or after the date they received a resident visa. In the year before they received a resident visa, roughly 1% of our sample found a new firm each month. In the months after, this rate increases dramatically, reaching a maximum 6 months later of 3.3%. The job-switching rate then begins to shrink, though it remains higher than the baseline rate.

The red series in Figure 2 Panel A depicts the job-switching rate for the worker's baseline colleagues. To construct it, we first find the average outcome among the baseline colleagues of each new resident, and then average over all new residents. The job-switching rate in the baseline colleague series begins at about 2.5% — higher than the corresponding rate for new residents because most of these colleagues are already residents or citizens. There is no increase in the baseline colleague series around the time when the modal worker becomes a resident. (The dip in the baseline month is due to our requirement that both the new resident and her controls have a unique firm in that month.)

Figure 2 Panel A suggests that Essential Skills job restrictions are binding: when these restrictions are removed, many workers switch firms. Figure 2 Panel B asks whether these restrictions reduce the worker's earnings. There is no evidence that they do. Earnings for the new resident follow a stable trend around the date they receive a resident visa. Baseline earnings for their colleagues are very similar before the modal worker receives a resident visa (despite us not matching on colleagues' earnings) and continue to be similar after.

Appendix Figure A1 repeats this analysis for four other variables: employment, visa status, weekly hours of work, and log firm size. New residents are somewhat more likely to be employed both before and after they receive their resident visa. About 65% of control workers hold a resident visa or citizenship a prior to the modal migrant receiving a resident visa; this proportion increases smoothly to about 85% three years later. Conditional on working, new residents typically work no more or fewer hours than they did prior to receiving a resident visa.[17]

Panel D of Appendix Figure A1 shows that new residents disproportionately move to larger firms. In differentiated firm models (such as that in Section 2), larger firms either pay greater wages or provide better amenities. Given that obtaining a resident visa is not associated with increased earnings, it appears that new residents are disproportionately moving into high-amenity firms.

To provide causal point estimates, we calculate a matched regression of the form

$$y_{i,t} - \bar{y}^c_{i,t} = \beta + \delta\left(y_{i,-1} - \bar{y}^c_{i,-1}\right) + e_{i,t}, \tag{2}$$

where $y_{i,t}$ is migrant i's outcome t months after they received a resident visa, $\bar{y}^c_{i,t}$ is the average

[17] Recall that hours of work are measured using survey data. As such, Appendix Figure A1 Panel C is formed using only those workers who happen to be surveyed in a given month.

outcome among their baseline colleagues, and β is the coefficient of interest. We estimate the equation on months 0–36 after receiving a resident visa. Standard errors are clustered at the level of the baseline firm.

Results are presented in the blue series of Figure 3. (The underlying estimates for that figure are available in Appendix Table A2.) The left-most bar aggregates across all occupations. We estimate a precise zero, and are able to rule out any effects on log earnings larger than 0.008 with 95% confidence. For comparison, recall from Table 3 that the average resident earns about 20% more than the average migrant.

Whereas the first bar in Figure 3 aggregates across all occupations, later bars are estimated on subsets of occupations. Specifically, we divide workers on the basis of their baseline occupation: their occupation in the month before they received a resident visa. We first divide workers by whether their baseline occupation is included on an Essential Skills in Demand list at the time they received a resident visa. (One might expect to see larger effects for occupations *not* on an Essential Skills in Demand list — as discussed in Section 3, migrants' job restrictions were looser in listed occupations.) We find no effects for either occupations which are on an Essential Skills in Demand list or occupations which are not.

Existing work has found that professional occupations exhibit more bargaining than manual occupations (Hall & Krueger, 2012; Caldwell & Harmon, 2019; Caldwell et al., 2024). We calculate our estimator separately by high-level occupation. We find no evidence that a resident visa increases earnings in those occupations where we would most expect to see an effect: managers and professionals.

We do, however, find a positive and significant effect for workers in 'community and personal services'. To assess whether this effect is driven by these workers bargaining higher wages at their baseline firms, we re-estimate Equation (2), replacing the dependent variable with the worker's colleagues' log earnings:

$$\text{colleagues' log earnings}_{i,t} = \frac{1}{\left|C_{i,t}\right|} \sum_{j \in C_{i,t}} \text{log earnings}_{j,t},$$

where $C_{i,t}$ is the set of workers who share a firm with worker i in period t:

$$C_{i,t} \equiv \left\{ j : f_{i,t} = f_{j,t} \right\} \setminus \{i\}.$$

When constructing this variable, we exclude workers (and colleagues) with multiple firms in a given month, and we exclude the first and last month of employment at a given firm. We construct this variable for both the new residents and for their baseline colleagues, and can thus reestimate Equation (2).

The resultant estimates are presented in the red series of Figure 3. The positive effect on workers in community and personal services appears to be driven by firm changes: these workers' colleagues' earnings increase by almost as much as their own. Indeed, we find that workers generally move into higher-paying firms when they obtain a resident visa, even though their own earnings do not tend to increase.

This analysis has shown that firm changes are important. However, we have not yet shown that bargaining is unimportant. Define a worker's *relative log earnings* as their own log earnings minus their colleagues' log earnings:

$$\text{relative log earnings}_{i,t} \equiv \text{log earnings}_{i,t} - \text{colleagues' log earnings}_{i,t}.$$

Under a bargaining equilibrium, expanding a worker's job options will increase the relative log earnings *of workers who stay at their original firm.* However, even under a bargaining equilibrium, the effect on workers who move firms is ambiguous: workers might move to firms for which they have such a strong preference that their bargaining power decreases. It thus remains to be shown that expanding a worker's job options has no effect on workers who stay at their original firm — this is the task of Subsection 5.3. Before we do so, we will first respond to a distinct concern: that workers are obtaining their resident visas in anticipation of idiosyncratic shocks.

5.2 Lottery estimates

To construct the estimates reported in the previous subsection, we match workers to their baseline colleagues. As such, those estimates are unbiased even if workers apply for a resident visa in anticipation of a future economic shock — provided that the shock affects all workers in their occupation-by-firm cell equally. One natural concern is that workers apply for a resident visa in anticipation of an *idiosyncratic* shock. For example, it might be the case that workers apply for a resident visa when they become ill, or when they fall out of favor with their boss. If this is the case, the matched-control specification may be biased.

To address this concern, this subsection studies the subset of Essential Skills migrants who enter a random lottery for a resident visa. Specifically, we study migrants who enter a lottery for either the Samoan Quota Resident Visa or the Pacific Access Category Resident Visa. Like all New Zealand resident visas, these visas allow their holders to stay in New Zealand indefinitely. Citizens of Samoa are eligible to enter the lottery for the Samoan Quota Resident Visa while citizens of Kiribati, Tuvalu, Tonga and Fiji are eligible to enter the lotteries for the Pacific Access Category Resident Visa. Entrants from different origin countries enter different lotteries; each year, one lottery is held for each origin country. Typical lottery entrants reside in their home country. However, some are already in New Zealand when the lottery is drawn; these are the entrants we study.[18]

[18] Existing work has studied how winning one of these lotteries affects those who had previously been living in

Analytic sample and balance tests. We include all lottery entrants who, at the time of the lottery, held an Essential Skills visa and were in New Zealand. We include both primary applicants and secondary applicants (*i.e.* dependants). We exclude a small number of workers who entered a given lottery more than once (*i.e.* as both a primary applicant and as a secondary applicant).

Appendix Table A3 describes our sample. Entrants typically earn less, are more often men, and are more likely to live in Auckland compared to the general population of Essential Skills migrants. We observe 6123 entries across 3465 workers; 186 of these workers are successful. Entrants have typically been in New Zealand for between 3 and 4 years. 60% of successful entrants have obtained a resident visa within 12 months of the lottery, and 80% have within 24 months. Only 10% of unsuccessful entrants have obtained a resident visa 24 months after the lottery.

To confirm that the lottery is indeed random, we estimate the equation

$$y_{i,l} = \beta \text{success}_{i,l} + \alpha_l + e_{i,l}, \tag{3}$$

where $y_{i,l}$ is a placebo outcome for migrant i in lottery l, β is the coefficient of interest, $\text{success}_{i,l}$ indicates whether migrant i was successful in lottery l, α_l is a lottery fixed effect, and $e_{i,l}$ is an orthogonal error. (Different lotteries have different probabilities of success. The lottery fixed effect ensures that we compare successful entrants only to unsuccessful entrants in the same lottery.) Unsuccessful entrants can reenter later lotteries, so we observe some workers in more than one lottery. We estimate the equation with ordinary least squares, and to account for the correlation between different observations of the same worker we cluster standard errors at the worker level.

Table 4 reports these placebo estimates. Conditional on their lottery, successful and unsuccessful entrants have similar ages and numbers of dependants, and they have been in New Zealand for a similar length of time. They are similarly likely to be male, to live in Auckland, and to live in a major city. They have similar baseline employment rates and earnings. They are somewhat more likely to have switched firms in the 6 months prior to the lottery; given that we see no other evidence of imbalance between the successful and unsuccessful entrants we view this difference as spurious. The p-value testing the hypothesis that there is no difference across any outcome is 0.349.

Results. In Figure 4 we report OLS estimates of the coefficients β_t from the equation

$$y_{i,l,t} = \beta_t \text{success}_{i,l} + \alpha_{l,t} + e_{i,l,t}, \tag{4}$$

where i is a migrant who entered lottery l, and t equals the number of months since lottery l was drawn. In Panel A the dependent variable is an indicator for having received a resident visa

their home country (Stillman, McKenzie, & Gibson, 2009; McKenzie, Stillman, & Gibson, 2010; Gibson, McKenzie, Rohorua, & Stillman, 2019).

t months after the lottery. Evidently, winning the lottery has a very large effect on the likelihood that a migrant receives a resident visa, reaching a maximum of about 70 percentage points one year after the lottery. Thereafter, resident visas for successful entrants begin to plateau while those for unsuccessful entrants begin to catch up.

In Figure 4 Panel B the dependent variable is an indicator for having worked for a new firm at any point during the t months since the lottery. Winning a lottery has a substantial effect on the likelihood that a migrant switches firms. In Figure 4 Panel C the dependent variable is log earnings in the tth month since the lottery. There is no evidence that winning a lottery increases a migrant's earnings.

To provide point estimates, we use the lottery as an instrument for whether a migrant holds a resident visa. Specifically, we estimate the linear system

$$\begin{aligned} y_{i,l,t} &= \beta \text{resident}_{i,l,t} + \alpha_l + \epsilon_{i,l,t}, \\ \text{resident}_{i,l,t} &= \delta \text{success}_{i,l} + \gamma_l + u_{i,l,t}, \\ \text{success}_{i,l} &\perp \epsilon_{i,l,t}. \end{aligned} \tag{5}$$

The key assumption is that the lottery only affects outcomes through its effect on migrants' visa status. Given the institutional structure we study, that assumption is plausible. In our baseline specification we estimate the system using outcomes 12–36 months after the lottery (which Figure 4 indicates is the period in which our instrument has the largest effect on visa status).

Results are presented in Table 5. Receiving a resident visa increases the monthly job-switching rate by about 1.5 percentage points. However we find no evidence that a resident visa affects migrants' earnings; we can rule out any effect on log earnings larger than 0.01. These results are robust to various alternative specifications which control for baseline outcomes, which expand the set of months for which we measure outcomes, which omit workers who held an Essential Skills in Demand list occupation at the time of the lottery (for whom we might expect smaller effects), and which change the functional form of the estimator.

Appendix Table A4 reestimates our instrumental variables system to study how receiving a resident visa affects other outcomes. One might worry that receiving a resident visa affects the probability that a worker is employed in New Zealand, introducing a selection bias into our regression estimates. The result in Column (1) suggests this is not the case: receiving a resident visa has no detectable effect on the likelihood that a migrant is employed in New Zealand (although, as shown in Column (2), this is because a positive effect on being in New Zealand is offset by a negative effect on being employed conditional on being in New Zealand). In Column (3) of Appendix Table A4 we repeat the exercise presented in Subsection 5.1, where we ask whether a resident visa induces workers to move into higher-paying firms. We find no evidence that it does.

Our preferred specification, reported in Column (2) of Table 5, implies that obtaining a resident visa *decreases* migrants' earnings by about 6%. While this effect is not significant at the 5% level, it is consistent with a major concern with our analysis thus far: new migrants may be moving to lower-paying, higher-amenity jobs. Such behavior is consistent with models in which wages are set through bargaining, and so our results thus far do not refute such models. In the following section, we study an unambiguous prediction of bargaining models: that increasing a worker's job options will increase the earnings of those who remain at their original job.

5.3 The effect of a resident visa among workers who do not move jobs

The results discussed above demonstrate that expanding a worker's job options does not, on average, increase their wages: though new residents often move jobs, their average earnings are unaffected. In bargaining models of wage-setting, expanding a worker's job options can affect their wages in two ways: the worker might move to one of their new job options, or the worker might use their improved outside job options to negotiate a higher wage at their existing job. Though the latter effect is unambigiously positive, the former effect could be either positive or negative. In concrete terms, new residents appear to be moving to higher-amenity jobs. If higher-amenity jobs tend to pay lower wages, a negative *job-switching effect* could be obscuring a positive *bargaining effect*.

Per Proposition 1, the ideal test of bargaining asks whether receiving a resident visa increases the earnings of those workers who choose not to move jobs. In this subsection, we estimate the 'stayers average treatment effect' using Subsection 5.2's sample of workers who enter a resident visa lottery.

Identification overview. Workers who are paid poorly at their current job will, if given the chance, be more likely to move jobs. As such, if we simply condition on workers who stay at their prior job, the visa lottery will no longer be independent of workers' potential outcomes. Nonetheless, under plausible assumptions, the stayers average treatment effect *can* be identified.

We observe the realized earnings of lottery-winning job stayers. The identification challenge is thus to identify *counterfactual* earnings for lottery-winning job-stayers: what they would have earned, on average, had they lost the lottery. To do so, we must infer how average counterfactual earnings among lottery-winning job-stayers differ from average earnings among lottery-losers. We consider three approaches, which require different sets of assumptions. Below, we sketch the identification arguments informally; formal identification proofs are provided in Appendix B.

Lee bounds. Our first approach imposes only the assumption that the decision to move jobs is monotonic in winning the lottery. Concretely, this assumption requires that lottery-winners who did not move jobs would still not have moved had they lost the lottery, and that lottery-losers who *did* move would still have moved had they *won* the lottery. This monotonicity assumption is plausible in our context given that winning the lottery expands the set of available job options. With this assumption, we can bound the stayers average treatment effect by dropping the highest or lowest outcomes among job-staying lottery losers. This approach is formally equivalent to the Lee (2009) bounds used by experimentalists to correct for differential sample selection.

Using the OLS estimate as an upper-bound. Our second approach imposes an additional assumption: that mean earnings for lottery-losing job-stayers are lower than mean counterfactual earnings for lottery-winning job-stayers. Concretely, this assumption requires that lottery-winners are disproportionately moving away from low-paying jobs. This assumption implies that the 'bad' OLS estimand, which simply conditions on the set of job-stayers, is unambiguously upwards-biased for the stayers average treatment effect. As such, we can use the OLS estimate as an upper bound.

We cannot test whether earnings among lottery-losing job-stayers tend to be lower than *counterfactual* earnings among lottery-winning job-stayers. However, we *can* test whether *baseline* earnings tend to be lower among job-staying lottery losers than among job-staying lottery winners. We find that they are, although the difference is not statistically significant.[19]

Modeling job-switching. In our third approach, we account for the selection bias that results from conditioning on job-stayers by modeling how a lottery-winner's decision to move jobs depends on their counterfactual earnings. We do not observe the relationship between a lottery-winner's job-moving and their *counterfactual* earnings. However, we do observe the relationship between a lottery-winner's job-moving and their baseline earnings. We also observe the relationship between lottery-losers' baseline earnings and lottery-losers' post-lottery earnings. Under plausible assumptions, these two observed relationships can be combined to identify the relationship between a lottery-winner's job-moving and their counterfactual earnings — and thus identify the stayers average treatment effect.

The key assumption is that job-moving is Markovian in counterfactual earnings: conditional on a worker's *current* counterfactual earnings, their decision to move jobs does not depend on their past counterfactual earnings. We view this assumption as plausible over short time frames when wealth effects, for example, are unlikely to be important.

This assumption yields nonparametric identification in a stylized setting under which lottery losers *never* move jobs. In practice, some lottery losers do move jobs (though they move less

[19] Among job-stayers, regressing baseline log earnings on an indicator for having won the lottery (controlling for lottery fixed effects) yields a coefficient of 0.012 [standard error: 0.038].

often than do lottery winners). We thus cannot nonparametrically identify the data generating process for counterfactual earnings.[20] We instead impose the following parametric structure.

Associate each worker i with a lottery $l(i)$. Consider a worker i and time t, where time is measured relative to the date of the worker's lottery (*i.e.* $t = 0$ is the quarter in which the lottery was held).[21] Let $Y_{i,t}(W,M)$ indicate their potential log earnings, where W indicates that the worker won their lottery and M indicates that the worker has moved firms between quarter 0 and quarter t. In turn, their decision to move firms will depend on whether they won the lottery, which we represent with another potential outcome $M_{i,t}(W)$. Letting W_i indicate whether the worker won actually their lottery, a worker's realized log earnings are thus $Y_{i,t} = Y_{i,t}\left(W_i, M_{i,t}\right) = Y_{i,t}\left(W_i, M_{i,t}(W_i)\right)$. The stayers average treatment effect is given by

$$\text{ATE}_{\text{stayers}} \equiv \mathbb{E}\left[Y_{i,t}(1,0) - Y_{i,t}(0,0)\,\middle|\,M_{i,t}(1) = M_{i,t}(0) = 0; t \in 4,\ldots,11\right].$$

To match the results reported in the previous subsection, we measure outcomes 4–11 quarters after the lottery. Note that $\text{ATE}_{\text{stayers}}$ measures the effect of winning the lottery, rather than the effect of receiving a resident visa. As such, it is most comparable to the 'intent-to-treat' estimates reported in Column (1) of Table 5, rather than the IV estimates reported in later columns.

To estimate $\text{ATE}_{\text{stayers}}$, we first condition on the population who have a consistent employer in the quarter prior to the lottery: $M_{i,-1} = 0$. Among this population, we impose the following nonparametric assumptions:

1. If a worker who lost the lottery moves, she would have moved had she won the lottery: $M_{i,t}(1) \geq M_{i,t}(0)$.

2. The lottery has no effect on earnings before it is held: $Y_{i,-1}(1,0) = Y_{i,-1}(0,0)$.

In our baseline specification, we additionally impose the following parametric assumptions:

3. The decision to move, among workers who have not done so already, is probit in W_i, $Y_{i,t}(0,0)$, and their interaction, with a lottery effect:

   $$\mathbb{P}\left[M_{i,t} = 1 \mid W_i, \left(Y_{i,s}(0,0)\right)_{s\leq t}, M_{i,t-1} = 0\right] = \Phi\left(\beta_0 + \beta_1 W_i + \beta_2 Y_{i,t}(0,0) + \beta_3 W_i Y_{i,t}(0,0) + \delta_{l(i),t}\right);$$

   where Φ is the normal CDF.

[20] This is a 'full support' issue, closely related to the 'identification at infinity' issue with identifying the Roy model (Taber & Vejlin, 2020).

[21] In our parametric model we measure outcomes at the quarterly level, rather than at the month level, to minimize transitory measurement error due, for example, to some months having more fortnightly pay cycles than others. This simplifies the DGP for counterfactual earnings that we must specify.

4. Counterfactual earnings $Y_{i,t}(0,0)$ are AR(1), with another lottery effect:

$$Y_{i,t}(0,0) = \alpha_0 + \alpha_1 Y_{i,t-1}(0,0) + \gamma_{l(i),t} + \epsilon_{i,t}; \quad \epsilon_{i,t} \sim N(0, \sigma_\epsilon).$$

5. Lottery effects are joint-normally distributed:

$$\left(\delta_{b,t}, \gamma_{b,t}\right) \sim N\left(0, \Sigma_{\delta\gamma}\right).$$

We calculate the Bayesian posterior of the above model, given uninformative priors, using a Gibbs sampler described in Appendix C.

Analytic sample. We here use the same sample of lottery entrants studied in Subsection 5.2. In our specification that parametrically models job-switching we additionally retain only workers who are observed in 3 months prior to the month in which the lottery was held and who did not move in those three months. In that specification we also collapse our data to the worker-by-lottery-by-quarter level (rather than the worker-by-lottery-by-month level at which we conduct our other analyses). As mentioned above, measuring earnings at the quarterly level reduces transitory measurement error.

Results. Estimates of the stayers average treatment effect are presented in Table 6. In columns (1) and (2) we report Lee bounds: the specification in Column (1) does not condition on any covariates while the specification in Column (2) conditions on 5 quintiles of baseline earnings. In both these specifications our Lee bounds span zero, though the confidence intervals are large and we cannot rule out economically meaningful effects.

In Column (3) of Table 6 we present the OLS upper-bound. This upper-bound is negative, and a 95% confidence interval excludes any effect on log earnings larger than 0.024.

In Column (4) of Table 6 we present an estimate that exploits our parametric model of job-switching. We again find no evidence that winning a resident visa lottery increases the earnings of job-stayers. Our point estimate is negative, and our posterior distribution places very low probability on economically significant positive effects.[22]

In the results presented in Column (5), we weaken our assumption that job-switching is Markovian in counterfactual earnings. Specifically, we add a lag of counterfactual earnings to the selection equation

$$\begin{aligned}
&\mathbb{P}\left[M_{i,t} = 1 \mid W_i, \left(Y_{i,s}(0,0)\right)_{s \le t}, W_{i,t-1} = 0\right] \\
=&\Phi\left(\beta_0 + \beta_1 W_i + \beta_2 Y_{i,t}(0,0) + \beta_3 W_i Y_{i,t}(0,0) + \beta_4 Y_{i,t-1}(0,0) + \beta_5 W_i Y_{i,t-1}(0,0) + \delta_{l(i),t}\right).
\end{aligned}$$

Again, we find no evidence that winning a resident visa lottery increases the earnings of those who stay at their original firms.

[22] Appendix Table A5 lists estimates of the model primitives. These are plausible: moving firms is positively associated with winning the lottery and is negatively associated with counterfactual earnings, with lottery-winners being particularly sensitive to their counterfactual earnings.

6 How Loosening Restrictions in an Occupation Affects Wages in that Occupation

In the previous section, we showed that expanding a worker's job options does not tend to increase their earnings. However, this *individual-level* result need not imply that a *market-level* shock to workers' job options will have no effect. As shown by Proposition 2, in a Bertrand wage-posting equilibrium, expanding the job options of many workers can increase both their wages and the wages of other workers. In this section, we test the effects of market-level shocks directly.

To do so, we study how wages in an occupation change when that occupation is added or removed from an Essential Skills in Demand list. As discussed in Section 3, in occupations included on one of these lists, migrants could work for any firm willing to employ them. We ask how expanding migrants' job options affects both their wages and the wages of their non-migrant colleagues.

Data. Over our sample period, the relevant Essential Skills in Demand lists were the 'Construction and Infrastructure Skill Shortage List', the 'Canterbury Skill Shortage List', the 'Immediate Skill Shortage List', the 'Long Term Skill Shortage List' and the 'Regional Skill Shortage List'. These lists differed mostly in the process used to amend them, in whether they only applied to certain sub-national regions, and whether they were also used to assess applications for a resident visa. We obtained historic lists from government archives and from the Internet Archive.

We treat an occupation as being included on a list in a given month if it was included in any list on the first of that month. Some lists only expanded migrants' job options in certain regions. To avoid *a priori* assumptions about migrants' willingness to move across regional boundaries, we consider an occupation treated if it was included for any region. We observe workers' occupations at the ANZSCO 6-digit level (*e.g.* 'Electrical Engineer'). Occasionally, the occupation included on an Essential Skills in Demand list is more specific (*e.g.* 'Power Systems Engineer'); in such cases we consider the entire ANZSCO 6-digit occupation as being treated.

We conduct our analysis using an occupation-by-month panel. The panel is described in Appendix Table A6. Occupations which were sometimes or always on an Essential Skills in Demand list had somewhat higher earnings than other occupations.

Identifying variation. We identify the effect of Essential Skills in Demand lists by asking how earnings change when an occupation is added or removed from these lists. The Immediate Skill Shortage List (which was later renamed as the Regional Skill Shortage List) and the Long Term Skill Shortage List were together amended in a single process. During the first years of our sample, these lists were supposed to be reviewed twice-yearly, though reviews were often delayed or postponed. During the later years of our sample, these lists were reviewed yearly. Amendments were proposed either by industry bodies or by the responsible government agency. These proposals were assessed internally, and final decision-making authority lay with a senior bureaucrat (for the Immediate Skill Shortage List) or with a government minister (for the Long Term Skill Shortage List).

In September 2011 — following the September 2010 and February 2011 earthquakes in the Canterbury region — the government introduced a special skill list for occupations required in the post-quake rebuild. This list initially included only occupations which were already included in other lists. In later iterations it included some occupations which were absent from other lists. This list was later renamed the Construction and Infrastructure Skill Shortage List, at which point it begun to include regions other than Canterbury. These lists were reviewed somewhat more regularly and informally than the lists described in the paragraph prior.

The criteria for including an occupation changed over time and varied between lists, but were most explicit for the Immediate Skill Shortage List in the latter half of our sample. During that period, proposals to amend the Immediate Skill Shortage List were assessed according to three criteria: 'skill level', 'scale', and 'shortage'. The 'skill level' criterion was satisfied provided the occupation was coded as sufficiently skill-intensive. The 'scale' criterion was satisfied provided the occupation either included at least 2000 workers, or had at least 50 Essential Skills visas approved in the prior 12 months. The 'shortage' criterion was assessed using various factors including (a) the number of unemployed 'jobseekers' registered with the New Zealand welfare agency, (b) the number of online job listings, (c) the number of recent visa approvals in that occupation, and (d) historic and forecast employment growth in the occupation.[23]

Some of these criteria are plausibly orthogonal to demand and supply shocks. For example, the 'scale' criterion was used to justify removing several small occupations during the later years of our sample, including upholterer, glider pilot instructor, rheumatologist, and geophysicist. However, the use of employment forecasts to assess the shortage criterion does threaten our differences-in-differences identification strategy: if occupations are listed in anticipation of future demand shocks, our estimated effect on earnings will be upwards-biased. We will assess this threat by using occupation-level Australian data and firm-level financial data as proxies for labor demand.

Results. We begin by estimating a linear panel event-study model (Freyaldenhoven, Hansen,

[23]The memorandum summarizing the December 2016 decision to remove 'upholsterer' from the Immediate Skill Shortage List is representative: "The Furniture and Cabinetmaking Asociation of NZ Inc. advised that with low cost imports coming into New Zealand, there are only a few small businesses now operating. There is little upholstery manufacturing going on in New Zealand, and apprentice numbers are low. They are quite relaxed about the removal of this occupation from a shortage list. There are small numbers of employees in this occupation and numbers of work visas granted over the last few years have been low. Occupation growth has been negative over the last five years and forecast growth is quite low. There are a small number of jobseekers registered with MSD who have claimed at least 12 months' work experience as an upholsterer."

Pérez, & Shapiro, 2021). That is — letting o index occupation, letting t index months, letting $\text{list}_{o,t}$ indicate inclusion on an Essential Skills in Demand list and letting Δ be the first difference operator — we estimate the linear equation

$$\text{mean log earnings}_{o,t} = \delta_{-13}\left(1 - \text{list}_{o,t+12}\right) + \sum_{s=-12}^{36} \delta_s \Delta \text{list}_{o,t-s} + \delta_{37}\text{list}_{o,t-37} + \alpha_o + \gamma_t + \epsilon_{o,t}$$

with ordinary least squares, subject to the constraint $\delta_{-1} = 0$.

We depict these estimates in Figure 5. Earnings are stable before an occupation is added to a skill list. They increase quickly after an occupation is added, stabilizing at about 3.5% higher 6 months after the occupation is added. This result is consistent with wages being set in a Bertrand wage-posting equilibrium: when an occupation is added to an Essential Skills in Demand list, migrants' job options expand. This increases the elasticity of their labor supply; firms respond by setting more generous wages.

The event study depicted in Figure 5 suggests that the effect of listing an occupation can be well-approximated by a model in which the effect is instantaneous and lasts indefinitely:

$$\text{mean log earnings}_{o,t} = \beta\, \text{list}_{o,t} + \alpha_o + \gamma_t + \epsilon_{o,t}, \tag{6}$$

where $\text{list}_{o,t}$ indicates that the occupation is included on an Essential Skills in Demand list, α_o is an occupation fixed effect, γ_t is a month fixed effect, and β is the parameter of interest. We estimate the two-way fixed effects model in Equation (6) with ordinary least squares. The estimated effect — depicted in Column (1) of Table 7 — is equal to about 0.03. This is consistent with the estimates of the linear panel event-study model.

Proposition 2 told us that expanding migrant's job options can increase *both* migrants' and residents' earnings. In Columns (2) and (3) of Table 7 we assess this prediction. We find that listing an occupation increases the earnings of both migrants and residents.

Our identification strategy will be threatened if occupations are added to an Essential Skills in Demand list in anticipation of positive labor demand shocks. We here assess this threat in two ways.[24] We first test whether our estimates are sensitive to the inclusion of controls. Column (4) of Table 7 reestimates our two-way fixed effect model, additionally controlling for log firm value-added and for the log number of workers at a firm. (We construct these variables at the worker level, and then take occupation averages.) Column (5) includes these controls and additionally controls for occupation-specific linear time trends. In both columns (4) and (5), the estimated effect is similar to that in our baseline specification.

[24] Under 'Mechanisms' below we estimate regressions with mean firm productivity as the dependent variable; doing so tests the joint hypothesis that additions to Essential Skills in Demand lists are orthogonal to firm productivity shocks *and* that additions to Essential Skills in Demand lists don't move workers into more productive firms.

We next assess whether occupations are listed in anticipation of positive labor demand shocks by studying a proxy for labor demand shocks: wages in Australia. It is unlikely that New Zealand work visa policy affects Australian wages.[25] Thus, were we to find that Australian wages increase when an occupation is included on a New Zealand Essential Skills in Demand list, we would conclude that occupations are being listed in anticipation of positive demand shocks.[26]

We operationalize this test using publicly available data on mean Australian earnings, by occupation, provided in Table 14B of the Australian Tax Office's Individuals Statistics release. This data is released at the 4-digit ANZSCO occupation level, for each Australian tax year. Recall our New Zealand data is available at the 6-digit ANZSCO-by-month level: For this comparison we aggregate the New Zealand data to the 4-digit ANZSCO-by-year level. We consider a 4-digit ANZSCO occupation as treated in a given year if any of its component 6-digit occupations are included on an Essential Skills in Demand list at any point during that year. The Australian data is available beginning from the July 2010–June 2011 tax year. In the Australian data we do not observe mean *log* earnings; we instead take the log of mean earnings.

Figure 6 presents linear panel event study estimates which ask how both New Zealand earnings and Australian earnings are affected by the inclusion of an occupation on a New Zealand Essential Skills in Demand list. As in the more granular specification presented in Figure 5, New Zealand earnings increase when an occupation is included on a New Zealand Essential Skills in Demand list. There is no similar effect on Australian earnings. (We do find weak evidence of a pre-trend in the Australian data. Given that this pretrend ends when an occupation is listed, and that we find no similar pre-trend in the New Zealand data, we interpret it as spurious.)

Columns (6) and (7) of Table 7 exploit the Australian data to estimate point estimates for the effect of the Essential Skills in Demand list on New Zealand earnings. Column (6) estimates a triple-difference specification:

$$\text{New Zealand mean log earnings}_{o,t} - \text{Australian log mean earnings}_{o,t} = \beta\, \text{list}_{o,t} + \alpha_o + \gamma_t + \epsilon_{o,t}.$$

Column (7) presents the Freyaldenhoven, Hansen, and Shapiro (2019) estimator, in which Australian log mean earnings are used as a proxy for an unobserved confound. Both approaches

[25] Though some workers do move from New Zealand to Australia, they are a small proportion of the Australian labor force. In 2012, 2.5% of Australian residents were born in New Zealand `https://web.archive.org/web/20240306215258/https://www.abs.gov.au/statistics/people/population/australias-population-country-birth/latest-release`.

[26] Wages in Australia and New Zealand tend to co-move, suggesting that the two countries face correlated labor demand shocks. An OLS estimate of the linear equation

$$\text{log mean Australia earnings}_{o,t} = \beta\, \text{mean log New Zealand earnings}_{o,t} + \alpha_o + \gamma_t + \epsilon_{o,t}$$

yields a slope coefficient of 0.218, with a standard error (allowing for within-occupation clustering) of 0.078.

result in point estimates similar to our baseline estimate.

Two-way fixed effects models — like that in Equation (6) — can perform poorly when treatment effects are heterogeneous (De Chaisemartin & d'Haultfoeuille, 2020). The Essential Skills in Demand lists will likely have greatest effect when many workers are migrants. As such, we now consider an estimator which recovers an average effect of inclusion on an Essential Skills in Demand list.

In Appendix D we define our estimator formally. The key insight is that a literal difference-in-differences — earnings growth in a newly-treated occupation minus earnings growth in a control occupation — *is* an unbiased estimator of the treatment effect in the newly-treated occupation if the treatment and control occupations had both been untreated for long enough that any effects of historical treatments had dissipated (De Chaisemartin & d'Haultfoeuille, 2020; De Chaisemartin & d'Haultfoeuille, 2024). We study events in which an occupation is either added to an Essential Skills in Demand list after being absent for at least 24 months or is removed after being included for at least 24 months. For each event we identify a set of control occupations with the same pre-event status. We calculate the simple difference-in-difference for each event-control pair, measuring outcomes up to 36 months after the event. We then average these estimates to construct a single point estimate.

This point estimate is presented in Column (8) of Table 7. It is very similar to our baseline estimate but is slightly more precise.

One convenient feature of the nonparametric average treatment effect estimator is that it can easily be decomposed into subsets of occupations. In Appendix Figure A2 we report the average effect, conditional on the proportion of workers in the treated occupation who hold an Essential Skills visa. The results are noisy, but we do generally find larger point estimates in occupations with a greater proportion of Essential Skills migrants. This relationship is robust to whether we population-weight occupations when calculating average effects.

Mechanisms. In Table 8 we ask how including an occupation in an Essential Skills in Demand list affects intermediate outcomes. In Column (1), the dependent variable measures the proportion of workers, in an occupation, employed at a firm for the first time during that month. We find including an occupation in an Essential Skills in Demand list increases job-switching by about 10% of the baseline rate. This is consistent with the Essential Skills in Demand lists increasing labor market competition.

Proposition 2 told us that, in a competitive equilibrium, expanding the set of firms eligible to employ migrants should have no effect on wages. This prediction would seem to be rejected by the results in this section, which has found that expanding the set of firms eligible to employ migrants increases wages. However, the hypothesis of Proposition 2 shut down two channels through which expanding the set of firms might affect wages: Proposition 2 both assumed that firms have a homogeneous production function and it assumed that this production function exhibits constant returns to scale. When either of these conditions fail, expanding migrants' job options *can* decrease competitive equilibrium wages. The remaining columns of Table 8 test these conditions directly.

Columns (2) and (3) of Table 8 ask whether our results can be explained by concavity in firms' production functions. In Column (2) the dependent variable is simply the log number of workers in that occupation during the month; in Column (3) the dependent variable is the mean of the log number of workers across each worker's firm. Using either dependent variable we find that including an occupation in an Essential Skills in Demand list *increases* the number of workers, as prospective migrants can more easily obtain work visas. As such, the positive effects we have found on workers' earnings cannot be explained by concavity in firms' production functions.

Columns (4) and (5) of Table 8 use firm-level financial data to assess whether our results can be explained either by the reallocation of workers to more productive firms or by coincident productivity shocks.[27] For both columns we first construct a firm-by-year productivity measure. We then form the occupation-by-month dependent variable by taking the mean (across all workers in that occupation during that month) of the productivity of each worker's employer during that month. In Column (4), the productivity measure is the log ratio of a firm's revenue to its expenditure on intermediate inputs (which is a proxy for TFP). In Column (5), the productivity measure is log value-added per worker, where value-added is the firm's revenue minus its expenditure on intermediate inputs.

Positive coefficients in columns (4) and (5) of Table 8 would suggest either that inclusion in an Essential Skills in Demand list reallocates workers to more productive firms or that occupations are being listed in anticipation of positive productivity shocks. We find no evidence of either force. The coefficient in Column (4) is a precise zero, suggesting no relationship between inclusion in an Essential Skills in Demand list and firm productivity. The coefficient in Column (5) is negative. This negative coefficient is consistent with firms' production functions exhibiting decreasing returns to scale, with the increase in firm sizes (reported in Column (3)) reducing workers' average productivity.

Including an occupation in an Essential Skills in Demand list appears to increase wages despite the fact that listing an occupation increases the number of workers in that occupation. In other words, any reduction in workers' marginal product is more than offset by the fact that expanding migrants' job options reduces firms' market power. In the following section, we will isolate the effects of migrants' job options by estimating a structural model of the labor market.

[27] The underlying data for these outcome variables comes from firm surveys and firm tax returns, compiled in the Longitudinal Business Database (Fabling & Maré, 2019).

7 Specifying and Estimating a Structural Wage-Posting Model

This paper asks how restricting migrants' job options affects wages, profits, and welfare. Our analysis of market-level shocks has shown that restricting migrants' job options reduces average wages. Our analysis of individual-level shocks suggests that non-pecuniary preferences are also important: when an individual migrant's job restrictions are lifted, they often move firms despite not, on average, receiving higher wages.

However, these reduced-form analyses have limitations. We do not observe individual workers' preferences over firms, and thus a reduced-form analysis cannot identify welfare effects. Moreover, restricting migrants' job options will certainly affect different workers differently. This treatment effect heterogeneity is not, without further assumptions, identified in a reduced-form analysis (Heckman, Smith, & Clements, 1997). Finally, though restricting migrants' job options reduces average wages, these restrictions might not increase average profits — in particular, firms which cannot access migrant labor are likely to suffer. Estimating profit effects using a reduced-form analysis of occupation-level shocks would be difficult given that firms typically employ many occupations.

In this section we overcome these issues by estimating a structural model of the labor market. Our estimation procedure will assume a wage-posting equilibrium, which is consistent with the results in previous sections. The model is similar to that presented in Section 2. To facilitate estimation we impose functional forms on production and preferences. We will estimate our model using panel data, so we add a time dimension to our model, but the model will remain essentially static in the sense that agents' choice sets are unaffected by historical actions.[28] We also add workers' occupations and productivity as explicit components of the model (as in the theoretical extension provided in Appendix A).

In Section 8 we will use the model to understand the equilibrium effects of the Essential Skills visa system and to ask how the Essential Skills visa system mediates the wage effects of migration. To that end, the model quantifies two mechanisms by which restricting migrants' job options could affect the wage distribution:

1. Restricting migrants' job options will decrease the elasticity of their labor supply. When firms set wages they trade off the desire to attract more workers against the desire to pay lower wages. This trade-off is weaker when workers' labor supply is less elastic, and so

[28] In particular, we model only the *static* restrictions imposed by the Essential Skills visa system — that Essential Skills migrants can work only at certain firms — and ignore the *dynamic* restrictions — that Essential Skills migrants who switch firms must obtain a new visa. Our sense is that this assumption is innocuous: the cost of obtaining a new visa was relatively low (typically about two days' wages, and often free). However insofar as dynamic restrictions are important, our structural estimates will understate the restrictiveness of the Essential Skills visa system.

restricting migrants' job options will encourage firms to pay lower wages.

2. Restricting migrants' job options will segregate them into certain firms. If production exhibits decreasing returns to scale, then increasing the number of workers employed at a firm will decrease their marginal product of labor, and so decrease wages. The excluded firms, at which migrants cannot work, will employ fewer workers. Wages at these excluded firms would thus increase — unless there are complimentarities between different occupations, in which case the wages of some occupations might also decrease.

Quantifying the first mechanism requires that we quantify how restricted migrants' job options are in practice. For example, if we tell a Wellington-based carpenter that he can only work for one Wellington construction firm — but that he can work for any construction firm in Christchurch — have we substantially reduced his job options? We estimate workers' willingness to move both across firms and across space.

Quantifying the second mechanism requires that we quantify how employment of one worker will affect the marginal product of other workers. For example, if we allow a construction firm to employ migrant carpenters, will the marginal product of unskilled laborers increase or decrease? We estimate production functions allowing for both decreasing returns to scale and for complementarities between different occupations.

7.1 The Structural Model

We estimate a partial-equilibrium perfect-information model. In each period t, the model comprises a continuum of workers $\mathbf{I}_t$ and a finite set of firms $\mathbf{F}_t$.

During each period t, each worker $i \in \mathbf{I}_t$ has an (exogenous) occupation $o_{i,t}$ and each firm $f \in \mathbf{F}_t$ utilizes a set of occupations $\mathbf{O}_{f,t}$. Let $\mathbf{F}_{o,t} \equiv \{f \in \mathbf{F}_t : o \in \mathbf{O}_{f,t}\}$ denote the set of firms which employ occupation-o workers during period t.

Workers' choice sets. In each period t, each worker i chooses a single firm $f_{i,t}$ at which she will be employed. The set of firms from which she can choose depends on her occupation and on her visa status. Specifically, each worker is either a *migrant* or a *resident*. Let $\mathbf{F}_{o,t}^{migrant} \subset \mathbf{F}_{o,t}$ denote the subset of firms at which migrants in occupation o can work in period t. Residents can work at any firm. Let $\mathbf{F}_{i,t} \in \left\{\mathbf{F}_{o_i,t}^{migrant}, \mathbf{F}_{o_i,t}\right\}$ denote the set of firms at which worker i can work. In each period t, let $\mathbf{I}_t^{migrant} \subset \mathbf{I}_t$ denote the subset of workers who are migrants and let $\mathbf{I}_t^{resident} \equiv \mathbf{I}_t \setminus \mathbf{I}_t^{migrant}$ denote the subset who are residents.

Production. During each period t, each worker $i \in \mathbf{I}_t$ supplies effective labor $l_{i,t}$ to firm $f_{i,t} \in \mathbf{F}_{i,t}$. The total effective labor supplied to each firm in each occupation integrates over the effective labor supplied by each worker:

$$L_{f,o,t} \equiv \int_{i:f_{i,t}=f,o_{i,t}=o} l_{i,t} di.$$

Each firm f combines their labor to produce output using a CES production function:

$$y_{f,t}\left(\left(L_{f,o,t}\right)_{o\in\mathbf{O}_{f,t}}\right) = \left(\sum_{o\in\mathbf{O}_{f,t}} e^{\phi_{f,o,t}} L_{f,o,t}^{\rho}\right)^{\frac{\nu}{\rho}}, \tag{7}$$

where $\phi_{f,o,t}$ measures the firm's labor-augmenting productivity in occupation o, ν measures returns to scale and ρ measures firms' ability to substitute across occupations.

Output is sold at unit price in a competitive market.

Workers' preferences. Workers' firm choices are determined by their preferences over earnings and over the non-wage amenities provided by each firm. Each firm f has an exogenous, time-invariant location $c_f \in \mathbf{C}$. Denote the set of firms employing occupation-o workers in location c as $\mathbf{F}_{c,o,t} \equiv \left\{f \in \mathbf{F}_{o,t} : c_f = c\right\}$. Each worker i has a time-invariant demographic type $x_i \in \mathbf{X}$. This demographic type is unaffected by and distinct from the worker's visa status. In our application, it indicates whether the worker was born in New Zealand.

We assume that each worker i's preferences across firms f and earnings W can be represented by this utility function:[29]

$$u_{i,t}\left(W,f\right) = \tau_{x_i} \log\left(W\right) + \frac{1}{\lambda_{x_i}}\left(\bar{\xi}_{c_f,o_i,x_i,t} + \zeta_{i,c_f,t}\right) + \xi_{f,o_i,x_i,t} + \epsilon_{i,f,t}, \tag{8}$$

where τ_x measures the relative importance of earnings for demographic-x workers, $\bar{\xi}_{c,o,x,t}$ measures the typical amenity-value of location c among demographic-x workers in occupation o, $\xi_{f,o,x,t}$ measures the typical amenity-value of employment at firm f among demographic-x workers in occupation o, $\zeta_{i,c,t}$ measures i's idiosyncratic preferences across locations c (the importance of which is measured by the parameter λ_x) and $\epsilon_{i,f,t}$ measures i's idiosyncratic preferences across firms.[30]

We assume that workers' idiosyncratic preferences, demographic-type and visa status are independent of their effective labor. We also assume that, within each period, idiosyncratic

[29]This utility function is a monotonic transformation of that assumed in Section 2. Such a transformation has no effect on the wage-posting equilibrium, though it would affect the bargaining equilibrium because the Nash bargaining solution is affected by nonlinear transformations of agents' payoffs.

[30]We normalize the within-location amenity values $\xi_{f,o,x,t}$ by requiring that, within each location-occupation-demographic-year c,o,x,t: $\log\sum_{f\in\mathbf{F}_{c,o,t}}\exp\xi_{f,o,x,t} = 0$. We normalize the between-market amenity values $\bar{\xi}_{c,o,x,t}$ by requiring that, within each occupation-demographic-period o,x,t: $\log\sum_{c\in\mathbf{C}}\exp\bar{\xi}_{c,o,x,t} = 0$.

preferences follow a nested extreme value distribution, where the nests are the locations $\mathbf{C}$.[31] We do not restrict the dependence of idiosyncratic preferences across time.

Equilibrium. In each period t, each firm $f \in \mathbf{F}_t$ chooses a vector of occupation-specific wages $\vec{w}_{f,t} \equiv \left(w_{f,o,t}\right)_{o\in\mathbf{O}_{f,t}}$. A worker i employed at firm f_i in occupation o_i will receive earnings $W_{i,t} = w_{f_i,o_i,t} l_{i,t}$. Given wages, each worker $i \in \mathbf{I}_t$ chooses the firm that will maximize their utility:

$$f_{i,t} \in \underset{f\in\mathbf{F}_{i,t}}{\arg\max}\, u_{i,t}\left(w_{f,o_i,t} l_{i,t}, f\right). \tag{9}$$

Wages are set in Bertrand competition. Let $\vec{w}_{-f,t}$ denote the wages chosen by firms other than firm f. Let $\vec{L}_{f,t}\left(\vec{w}_{f,t}, \vec{w}_{-f,t}\right)$ denote the labor vector supplied to firm f, given the wages $\vec{w}_{f,t}, \vec{w}_{-f,t}$. Equilibrium wages will thus satisfy

$$\vec{w}_{f,t} \in \underset{\vec{w}\in\mathbb{R}^{|\mathbf{O}_{f,t}|}}{\arg\max}\left\{y_{f,t}\left(\vec{L}_{f,t}\left(\vec{w}, \vec{w}_{-f,t}\right)\right) - \vec{w}'\cdot\vec{L}_{f,t}\left(\vec{w}, \vec{w}_{-f,t}\right)\right\}.$$

7.2 Characterizing Equilibrium.

We will now characterize equilibrium wages and labor supply. We will use this characterization to estimate our structural parameters and to simulate counterfactuals.

Wages. Equilibrium wages will satisfy the first order conditions

$$\frac{\partial L_{f,o,t}}{\partial w_{f,o,t}}\left(\frac{\partial y_{f,t}}{\partial L_{f,o,t}} - w_{f,o,t}\right) - L_{f,o,t} = 0. \tag{10}$$

Rearranging Equation (10) yields the markdown equation

$$w_{f,o,t} = \frac{\partial y_{f,t}}{\partial L_{f,o,t}}\left(1 + \frac{1}{\eta_{f,o,t}}\right)^{-1}, \tag{11}$$

where $\eta_{f,o,t}$ is the period-t elasticity of occupation-o labor supplied to firm f.

Differentiating Equation (7) yields the marginal product of labor

$$\frac{\partial y_{f,t}}{\partial L_{f,o,t}} = \nu\left(\sum_{o'\in\mathbf{O}_{f,t}} e^{\phi_{f,o',t}} L^{\rho}_{f,o',t}\right)^{\frac{\nu-\rho}{\rho}} e^{\phi_{f,o,t}} L^{\rho-1}_{f,o,t}. \tag{12}$$

In Appendix E.2 we show that equations (11) and (12) can be combined to yield the log wage equation

$$\log w_{f,o,t} = \frac{\rho\log\nu}{\nu} + (\rho-1)\log L_{f,o,t} + \left(\frac{\nu-\rho}{\nu}\right)\log\tilde{L}_{f,t} - \log\left(1 + \frac{1}{\eta_{f,o,t}}\right) + \phi_{f,o,t}, \tag{13}$$

[31] That is, among demographic-x workers, they have CDF

$$\mathbb{P}\left[\left(\zeta_{i,c_f,t} + \lambda_x \epsilon_{i,f,t}\right)_{f\in\mathbf{F}_{o_{i,t},t}} \le \left(e_j\right)_{f\in\mathbf{F}_{o_{i,t},t}}\right] = \exp\left(\sum_{c\in\mathbf{C}}\left(\sum_{f\in\mathbf{F}_{c,o,t}}\exp\left(\frac{-e_f}{\lambda_x}\right)\right)^{\lambda_x}\right).$$

where $\tilde{L}_{f,t}$ is a measure of firm-level aggregate labor utilization given by

$$\tilde{L}_{f,t} \equiv \sum_{o \in \mathbf{O}_{f,t}} wL_{f,o,t}\left(1 + \frac{1}{\eta_{f,o,t}}\right), \tag{14}$$

and $wL_{f,o,t} \equiv w_{f,o,t}L_{f,o,t}$ is the total wage bill of occupation-o workers at firm f during period t.

Labor supply. Given the utility function (8), a worker's choice of firm will be unaffected by her effective labor:

$$\underset{f \in \mathbf{F}_{i,t}}{\arg\max}\; u_{i,t}\left(w_{f,o_i,t} l_{i,t}, f\right) = \underset{f \in \mathbf{F}_{i,t}}{\arg\max}\; u_{i,t}\left(w_{f,o_i,t}, f\right).$$

It follows that, conditional on demographic type and visa status, labor supply follows a standard nested-logit model. For occupation o, firm f and period t, define the mean utility of demographic-x residents as

$$\delta^{resident}_{f,o,x,t} = \tau_x \log(w_{f,o,t}) + \xi_{f,o,x,t}, \tag{15}$$

and define migrants' mean utility as

$$\delta^{migrant}_{f,o,x,t} = \begin{cases} \tau_x \log(w_{f,o,t}) + \xi_{f,o,x,t} & \text{if } f \in \mathbf{F}^{migrant}_{o,t} \\ -\infty & \text{otherwise} \end{cases}.$$

For either $status \in \{resident, migrant\}$, define mean utility for occupation-o demographic-x workers in location c as

$$\delta^{status}_{c,o,x,t} = \lambda_x \log\left(\sum_{f \in \mathbf{F}_{c,o,t}} \exp\left(\delta^{status}_{f,o,x,t}\right)\right) + \bar{\xi}_{c,o,x,t}. \tag{16}$$

The probability that a demographic-x occupation-o worker with $status \in \{resident, migrant\}$ will be employed at firm f in location c during period t, conditional on being employed by any firm within c, is given by:

$$\sigma^{status}_{f,o,x,t|c} \equiv \mathbb{P}\left[f_{i,t} = f \middle| i \in \mathbf{I}^{status}_t, o_{i,t} = o, x_i = x, c_{f_{i,t}} = c\right] = \frac{\exp \delta^{status}_{f,o,x,t}}{\sum_{f' \in \mathbf{F}_{c,o,t}} \exp \delta^{status}_{f',o,x,t}}. \tag{17}$$

Similarly, the probability that a demographic-x occupation-o worker with $status \in \{resident, migrant\}$ will be employed at any firm in location c is given by:

$$\sigma^{status}_{c,o,x,t} \equiv \mathbb{P}\left[c_{f_{i,t}} = c \middle| i \in \mathbf{I}^{status}_t, o_{i,t} = o, x_i = x\right] = \frac{\exp \delta^{status}_{c,o,x,t}}{\sum_{c' \in \mathbf{C}} \exp \delta^{status}_{c',o,x,t}}. \tag{18}$$

Let $\sigma^{status}_{f,o,x,t} = \sigma^{status}_{f,o,x,t|c_f}\sigma^{status}_{c_f,o,x,t}$ denote the overall proportion of demographic-x occupation-o workers with $status \in \{resident, migrant\}$ employed at firm f during period t. The supply elasticity of such workers to firm f will be

$$\eta^{status}_{f,o,x,t} = \tau\left(1 - (1-\lambda)\sigma^{status}_{f,o,x,t|c_f} - \lambda\sigma^{status}_{f,o,x,t}\right). \tag{19}$$

The overall supply elasticity to firm f aggregates over visa status and types:

$$\eta_{f,o,t} = \sum_{x \in \mathbf{X}} \left(\sigma_{x,resident|f,o,t} \eta^{resident}_{f,o,x,t} + \sigma_{x,migrant|f,o,t} \eta^{migrant}_{f,o,x,t} \right), \tag{20}$$

where $\sigma_{x,resident|f,o}$ is the proportion of firm f's occupation-o period-t workers who are demographic-x residents and $\sigma_{x,migrant|f,o}$ is the proportion of firm f's occupation-o period-t workers who are demographic-x migrants.[32]

7.3 Estimation

In this subsection we specify our estimation procedure for the parameter vector $(\tau_x)_{x\in\mathbf{X}}, (\lambda_x)_{x\in\mathbf{X}}$, ρ, ν. We identify these parameters under the assumption that shocks to the amenity value of employment are orthogonal to both contemporaneous productivity shocks and lagged productivity shocks. Specifically, to identify the labor supply parameters $(\tau_x)_{x\in\mathbf{X}}, (\lambda_x)_{x\in\mathbf{X}}$ we construct demand-shifting instruments using productivity shocks. Similarly, to identify the production function parameters ρ, ν we construct supply-shifting instruments using shocks to the amenity value of employment.

Our identification strategy is similar to that sometimes used to identify product market demand (MacKay & Miller, 2025). It requires that the non-wage reasons why a worker might choose to work at a particular firm — *e.g.* cultural fit, ease of public transport access, or amenities in the surrounding neighbourhood — change in a manner that is independent of both coincident and historic shocks to the firm's productivity. We view this assumption as plausible in the New Zealand context, where employment amenities rarely impose a literal cost on the employer. For example, employer-sponsored health insurance is rare.

Sample construction. We calculate both amenity shocks and productivity shocks using wages. However, we do not observe workers' effective labor $l_{i,t}$, and so wages will be measured with error. Under our assumption that effective labor is orthogonal to preferences and types, this measurement error will be classical. Nonetheless, it would bias estimation by introducing a spurious negative correlation between amenity shocks and productivity shocks.

We account for measurement error by randomly allocating all workers to one of two samples. We calculate the sample-s estimate of the log wage $\log \hat{w}^s_{f,o,t}$ by taking the mean log wage of workers in sample s. The wage-bill $wL_{f,o,t}$ can be observed without observing each worker's effective labor, and so we can treat $wL_{f,o,t}$ as being measured without error. We calculate the sample-s estimate of effective labor utilization $\hat{L}^s_{f,o,t}$ by dividing the wage-bill by the sample-s wage estimate: $\hat{L}^s_{f,o,t} = \frac{wL_{f,o,t}}{\hat{w}^s_{f,o,t}}$.

[32] Under our assumption that workers' effective labor is independent of their preferences and their type, labor supply elasticities equal *effective* labor supply elasticities .

We thus estimate our model using firm-occupation-year-demographic-sample-level data. A worker's demographic type indicates whether they were born in New Zealand. We exclude a small number of firm-occupation-year observations with a national market share of 1, for which our model implies an infinite markdown. We also exclude firm-occupation-year observations for which the national minimum wage is binding because the first order condition (10) need not hold for such observations. See Appendix E.1 for a complete description of this data.

Our fixed-point estimation procedure. Our estimator has a GMM representation — and indeed we use that representation (provided in Appendix E.3) to conduct inference. Nonetheless, we prefer the following 'fixed-point' representation because it makes clear the role of our identifying assumptions and because it is somewhat easier to calculate.

Given a guess for the labor supply parameters $\left(\tau_x^0\right)_{x\in\mathbf{X}}$, $\left(\lambda_x^0\right)_{x\in\mathbf{X}}$, consider the following routine:

1. Calculate the demographic-by-visa status labor supply elasticities $\eta_{f,o,x,t}^{status}$ using Equation (19), given observed market shares.

2. Calculate the overall labor supply elasticities $\eta_{f,o,t}$ by aggregating over visa statuses and demographic types, given the observed mix of workers at each job, using Equation (20).

3. Using these labor supply elasticities, calculate firm-level aggregate labor $\tilde{L}_{f,t}$ using Equation (14).

4. Using the observed labor supply of residents, invert equations (15) and (17) to calculate the amenity values $\xi_{f,o,x,t}$ (Berry, 1994). Do this separately for each sample's wage estimate $\hat{w}_{f,o,t}^s$; let $\hat{\xi}_{f,o,x,t}^s$ denote the amenity value calculated using the sample-s wage measure.[33]

5. Estimate the production parameters $\hat{\rho}, \hat{\nu}$. To do so, rearrange Equation (13) and take first-differences to form the estimating equation

$$\Delta \log \hat{w}_{f,o,t}^s + \Delta \log\left(1+\frac{1}{\eta_{f,o,t}}\right) = \left(\hat{\rho}-1\right)\Delta \log \hat{L}_{f,t}^s + \left(\frac{\hat{\nu}-\hat{\rho}}{\hat{\nu}}\right)\Delta \log \tilde{L}_{f,o,t} + \Delta\phi_{f,o,t}, \tag{21}$$

   where Δ is the first-difference operator. We estimate Equation (21) using linear IV, controlling for a sample-demographic-location-occupation-year fixed effect.

   As an instrument for $\Delta \log \hat{L}_{f,t}^s$ we use shocks to the amenity value of employment, calculating using the *other* split sample: $\Delta\hat{\xi}_{f,o,x,t}^{-s}$.

[33] An explicit formula for $\hat{\xi}_{f,o,x,t}^s$ is given in Appendix E.2.

We also require an instrument for firm-level aggregate labor $\Delta\tilde{L}_{f,t} = \Delta\sum_{o\in\mathbf{O}_{f,t}} e^{\mathcal{L}_{f,o,t}}$, where the summand $\mathcal{L}_{f,o,t}$ is given by $\mathcal{L}_{f,o,t} \equiv \log\left(wL_{f,o,t}\left(1+\frac{1}{\eta_{f,o,t}}\right)\right)$. To form that instrument, we first project the summands $\mathcal{L}_{f,o,t}$ on amenities (measured using the other sample) $\hat{\xi}^{s}_{f,o,x,t}$ to form $\mathcal{L}^{*-s}_{f,o,x,t}$. We instrument for $\Delta\log\tilde{L}_{f,o,t}$ using $\Delta\sum_{o\in\mathbf{O}_{f,t}} e^{\mathcal{L}^{*-s}_{f,o,x,t}}$.

6. Use these estimated parameters $\hat{\rho},\hat{\nu}$ to calculate labor-augmenting productivity $\phi_{f,o,t}$ by rearranging Equation (13). Let $\hat{\phi}^{s}_{f,o,t}$ denote the value of $\phi_{f,o,t}$ estimated using the sample-s wage estimate $\hat{w}^{s}_{f,o,t}$.

7. For each demographic type x, estimate the job-level labor supply elasticity $\hat{\tau}_x$. To do so, combine the equations governing residents' labor supply (15) and (17), take logs, and then take first-differences to form the estimating equation

$$\Delta\log\sigma^{resident}_{f,o,x,t|c_f} = \hat{\tau}_x\Delta\log\hat{w}^{s}_{f,o,t} + C_{s,c_f,o,x,t} + \Delta\xi_{f,o,x,t}, \tag{22}$$

where $C_{s,c,o,x,t}$ is a sample-location-occupation-demographic-year fixed effect. We estimate Equation (22) using linear IV. As an instrument for $\Delta\log\hat{w}^{s}_{f,o,t}$ we use *lagged* productivity shocks $\Delta\phi_{f,o,t-1}$. This instrument will be relevant provided that labor-augmenting productivity is mean-reverting.

8. For each demographic type x, estimate the market-level labor supply elasticity $\hat{\lambda}_x$. To do so, combine the equations governing residents' labor supply (15), (16) and (18), take logs, and then take first-differences to form the estimating equation

$$\Delta\log\sigma^{resident}_{c,o,x,t} = \hat{\lambda}_x\Delta\log\left(\sum_{f\in\mathbf{F}_{c,o,t}}\exp\left(\tau^0_x\log\hat{w}^{s}_{f,o,t} + \hat{\xi}^{s}_{f,o,x,t}\right)\right) + C_{s,o,x,t} + \Delta\bar{\xi}_{c,o,x,t}. \tag{23}$$

We estimate Equation (23) using linear IV. To form an instrument for $\Delta\log\left(\sum_{f\in\mathbf{F}_{c,o,t}}\exp\left(\tau^0_x\log\hat{w}^{s}_{f,o,t} + \hat{\xi}^{s}_{f,o,x,t}\right)\right)$, we project wages $\log\hat{w}^{s}_{f,o,t}$ on labor-augmenting productivity measured using the other sample $\hat{\phi}^{-s}_{f,o,t}$ to form $\log w^{*-s}_{f,o,t}$. We use $\Delta\log\left(\sum_{f\in\mathbf{F}_{c,o,t}}\exp\left(\tau^0_x\log w^{*-s}_{f,o,t}\right)\right)$ as our instrument.[34]

Our estimate of the labor supply parameters are the values that yield $\left(\tau^0_x\right)_{x\in\mathbf{X}} = \left(\hat{\tau}_x\right)_{x\in\mathbf{X}}$ and $\left(\lambda^0_x\right)_{x\in\mathbf{X}} = \left(\hat{\lambda}_x\right)_{x\in\mathbf{X}}$. Our estimate of the production parameters ρ,ν are the values given by Step 5 of the above routine, evaluated at our estimated labor supply parameters.

[34] To ensure our procedure is numerically identical to the GMM procedure given in Appendix E, in which we define moment conditions at the firm-occupation-year-demographic-sample level, when estimating Equation (23) we weight location-occupation-year-demographic-sample observations by the number of associated firm-occupation-year-demographic-sample observations.

Results. In Figure 7 we illustrate our fixed-point estimation procedure. We there consider a restricted estimator in which each of the labor supply parameters τ, λ has a constant value across between New Zealand-born and foreign-born workers. At each point in the x-axis of Figure 7 we consider a different candidate value of τ_0. (Across Figure 7 we fix $\lambda_0 = 0.131$, its estimate.)

As shown in Figure 7 Panel A, low values of τ_0 imply high production parameter estimates $\hat{\nu}$ and $\hat{\rho}$: a low value of τ_0 implies that most of the variation in employee growth is exogenous to productivity shocks, and thus the fact that there is only a weak relationship between employee growth and wage growth implies (per Equation (21)) that $\hat{\rho}, \hat{\nu} \approx 1$. However, these high values of $\hat{\rho}, \hat{\nu}$ imply (per Equation (13)) that wage growth is largely due to productivity growth and so is orthogonal to non-wage determinants of employee growth. The fact that employee growth is correlated with *lagged* wage growth thus implies a large value of $\hat{\tau}$ (as shown in Panel B). We can thus reject low values of τ.

Higher values of τ_0 imply lower values of $\hat{\nu}$ and $\hat{\rho}$, which in turn imply (modestly) lower values $\hat{\tau}$. The fixed point, at which $\hat{\tau} = \tau_0$, is depicted with dashed lines in Figure 7 and is also presented in Column (1) of Table 9. We find $\hat{\tau} = 0.814$. This parameter — which equals the labor supply elasticity facing an atomistic firm — is towards the lower end of labor supply elasticities estimated in the existing literature (Sokolova & Sorensen, 2021), though is higher than some experimental estimates (Dube, Jacobs, Naidu, & Suri, 2020). The low labor supply elasticity we find is consistent with New Zealand's relatively low labor share (Allan & Maré, 2021).

Our estimate of the market-level labor supply parameter $\hat{\lambda}$ is determined by a similar fixed-point requirement that $\hat{\lambda} = \lambda_0$. Our point estimate is $\hat{\lambda} = 0.131$, implying that the elasticity of migration to an atomistic *location* is $0.814 \times 0.131 = 0.107$: workers are largely unwilling to move to locations find a higher-paying job.

The resultant estimate of the returns-to-scale parameter is $\hat{\nu} = 0.940$. This parameter implies that firms face near-constant returns to scale: per Equation (12), increasing the number of employees (in every occupation) at a firm by 10% would decrease a worker's marginal product by only $10 \times (1 - \hat{\nu})\% = 0.6\%$. Firms can substitute across occupations easily. Our estimate of $\hat{\rho} = 0.919$ implies that the elasticity of substitution across occupations is $1/(1 - \hat{\rho}) = 12.3$.

In Column (2) of Table 9 we present our estimates that allow the labor supply parameters τ_x, λ_x to vary across demographic types. Existing work has hypothesized that migrants have less elastic labor supply, and that thus an increase in immigration could induce firms to set lower wages (Amior & Manning, 2021). Migrants' lower labor supply elasticity might, in part, be due to restrictions on their job options — the focus of this study — but it might also be due to other reasons, such as their having smaller referral networks. If migrants have less elastic labor supply only because their job options are restricted, then the negative wage effects of immigration could be avoided by reforming work visa policy. If migrants have less elastic labor supply for other reasons, the negative wage effects of immigration could be harder to avoid.

We find that, even among residents, foreign-born workers' labor supply is less elastic than that of New Zealanders. The difference is significant but relatively modest: we estimate $\hat{\tau}_{NZ} = 0.845$ and $\hat{\tau}_{foreign} = 0.751$. Per equations (11), (19) and (20), an atomistic firm that employed only New Zealand-born residents would, relative to an equally-productive firm that employed only foreign-born residents, pay a wage about $\left(1 + \frac{1}{\hat{\tau}_{foreign}}\right) / \left(1 + \frac{1}{\hat{\tau}_{NZ}}\right) - 1 \approx 7\%$ higher. Moreover, we find that foreign-born residents' location choices are *more* elastic than those of New Zealand-born residents: $\hat{\lambda}_{NZ} = 0.118$ and $\hat{\lambda}_{foreign} = 0.153$. As such, at firms that employ a large proportion of their local workforce, increasing immigration could induce *higher* wages. We return to the wage effects of immigration in Subsection 8.3 with equilibrium simulations that additionally account for firms' concave production functions.

8 Equilibrium Effects and Policy Counterfactuals

In this section we use the structural model — estimated in Section 7 — to understand both the overall effects of the Essential Skills visa system and how these effects might be mitigated. For simplicity, we study only the population at the 2018 census, and we drop time subscripts from notation. We first describe how we calculate equilibria, and then present the equilibrium effects of the Essential Skills visa system on wages, profits and output. We then use our model to ask how the Essential Skills visa system mediates the wage effects of increased immigration.

8.1 Calculating equilibria

Data. We use the same data with which we estimated our model in the previous section: we exclude occupation-year observations with a national market share of 1 (for which our model implies an infinite markdown); we exclude firm-occupation-year observations for which the national minimum wage is binding (for these observations the first order condition (10) need not hold); and we retain only observations with nonmissing wages in both samples. We calculate the model using the parameter vector reported in Column (2) of Table 9.

We calculate labor-augmenting productivity by inverting Equation 13. For firms or markets in which we observe type-x workers, we also calculate amenities using the equations derived in Appendix E.2. When we observe no type-x residents, we impute amenities using an algorithm described in Appendix E.4.

Simulating an equilibrium requires that we specify the set of firms which can employ migrants in each occupation.[35] We assume that a location-c firm can employ occupation-o migrants if either that firm was observed employing occupation-o migrants or occupation o was included in an Essential Skills in Demand list for location c.

Solving for an equilibrium. We solve for an equilibrium treating labor as continuous. We find the equilibrium wage vector by iteration: given an initial wage vector $\log w_r$, we find the mass of workers employed at each firm, and then find the implied wage vector $\log w'_r$ using equations (11) and (12). We then form the next wage vector by taking a convex combination of the initial wage vector and the implied wage vector: $\log w_{r+1} = 0.5 \times \log w_r + 0.5 \times \log w'_r$. We iterate this algorithm until convergence. We initialize this algorithm at the observed wage vector when finding the initial equilibrium and at the initial equilibrium when finding counterfactual equilibria.[36]

Our labor supply system is sufficiently flexible that our model could have multiple equilibria. We have found that starting the above fixed point routine from alternative starting values leads to a virtually identical equilibrium, which is consistent with there being a unique equilibrium at our estimated parameter vector.

Calculating worker-level effects. We are interested in both average effects and the distribution of effects. To calculate the latter, we allocate workers discretely across firms, given both an initial equilibrium wage vector and a counterfactual equilibrium wage vector. We do so using the algorithm presented by Townsend (in press).

8.2 How the Essential Skills visa system affects wages, profits and welfare

In our primary counterfactual, we allow Essential Skills migrants to work for any firm: $\forall o \in \mathbf{O}$: $\mathbf{F}_o^{migrant} = \mathbf{F}_o$. In Figure 8 Panel A, we depict how this counterfactual affects mean log wages.[37] Specifically, we depict the mean effect by both a worker's own visa status and by the proportion of her colleagues who themselves hold an Essential Skills visa. Regardless of their own visa status, workers who work with many migrants would experience a large increase in wages, were migrants' job options unrestricted. In contrast, migrants who work with few other migrants would actually tend to receive slightly lower wages were their job options unrestricted, because they tend to move toward lower-paying, higher-amenity firms.

[35] This set did not need to be specified for estimation — market shares sufficed.

[36] The observed wage vector need not be an equilibrium of our model: we sometimes impute amenities and, when calculating our equilibrium but not when calculating labor-augmenting productivity, we assume that workers have homogeneous effective labor. Across firm-occupation observations, the correlation between observed wages and wages in our initial equilibrium is 0.985. The correlation between observed worker counts and the mass of employed labor is 0.999.

[37] Occupations in which migrants' job choices are already unrestricted, because they are included on a nationwide Essential Skills in Demand List, are excluded from the population used to generate Figure 8 and Table A7.

On average, expanding migrants' job options would increase their equilibrium wage by about 5.6%. As discussed in Section 3, the typical resident works with few migrants, both because migrants are segregated into different firms and because relatively few workers are migrants. As such, the average worker is unaffected by the expansion of migrants' job options.

Figure 8 also depicts the distribution of effects across migrants (in Panel B) and residents (in Panel C). While many migrants receive a large wage increase, a substantial minority prefer a lower-paying job. The vast majority of residents are essentially unaffected. About 0.3% of residents would see their wage decrease by more than 2%, while 2.5% would see their wage increase by more than 2%. About 0.4% of residents would receive a wage increase greater than 10%.

In Appendix Table A7 we describe the subpopulations of workers represented in each bar of the Figure 8 histograms. Migrants who would have a lower wage, were migrants' job options unrestricted, typically have a higher baseline wage. Similarly, residents who would receive a log wage increase of at least 0.02 have an average log wage of 3.80 — much less than the average log wage among all residents, which is 3.92. Workers who would receive a large wage increase typically worked with more migrants at baseline.

The final column of Appendix Table A7 depicts the proportion of workers whose utility is greater in the counterfactual equilibrium than in the initial equilibrium. Expanding migrants' job options benefits almost all migrants: though many migrants have a substantially lower wage in the counterfactual equilibrium, these migrants have always moved to a new firm, and their overall utility is higher. Residents who receive a higher wage in the counterfactual equilibrium almost always prefer the counterfactual equilibrium. In addition, some residents who receive a lower wage in the counterfactual equilibrium are nonetheless better off: these are residents who are enticed to move to a lower-paying firm, because that firm pays more than it did initially.

Effects on profit and output. In Figure 9 we depict the effects on firm profit (in Panel A) and output (in Panel B). Both profits and output increase in firms which initially employ few migrants, and decrease in firms which initially employ many migrants. The overall effect is negative for aggregate profit and positive for aggregate output. Average effects are close to zero, because the large negative effects on the few firms which employ many migrants cancel out the small positive effects on the many firms which employ few migrants. Nonetheless, expanding migrants' job options would decrease aggregate profits by about \$116m, and increase aggregate output by about \$19m.

In Appendix Table A8 we compare the characteristics of firms which are benefited or harmed by the expansion in migrants' job options. Firms which benefit are typically smaller and less productive. They rarely employed many migrants in the initial equilibrium. Firms which are substantially harmed almost always employed migrants in the initial equilibrium, though there is an interesting group of firms who are slightly worse off despite employing no migrants initially: these are firms who lose their resident workers due to other firms' wage increases (or who *would* lose their residents, were they not to counter with their own wage increases).

The hetereogeneity of profit effects might explain the somewhat complicated politics of migrant job restrictions. In 2019 the New Zealand government proposed a mandatory accreditation scheme for employers of temporary migrants. While the scheme aimed to reduce extreme cases of migrant exploitation, it was generally seen as making it more difficult for employers to hire migrants. In consultation, 62% of employers supported the scheme (in whole or in part) while 25% were opposed.[38] Similarly, in 2018 the New Zealand government changed the Post Study Work Visa to allow these migrants to switch employers. In consultation, 41% of employers supported the proposal while 59% were opposed.[39]

Effects on welfare. We now calculate the effect of the Essential Skills visa restrictions on overall welfare. Let $w_{o,f}$ denote the initial wage paid by firm f to occupation o and let $w'_{o,f}$ denote the corresponding counterfactual wage. Let f_i denote worker i's initial employer, and let f'_i denote her counterfactual employer. We measure a worker i's 'money-metric' welfare WTP_i as their willingness to pay for their counterfactual employer and earnings:

$$u_i\left(w'_{o,f'_i} - WTP_i, f'_i\right) = u_i\left(w_{o,f_i}, f_i\right).$$

Given the form of the utility function (8), this equation has solution

$$WTP_i = w'_{o,f'_i} - w_{o,f_i}\exp\left(\frac{v_i - v'_i}{\tau}\right), \tag{24}$$

where $v_i \equiv \frac{1}{\lambda_{x_i}}\left(\bar{\xi}_{c_{f_i},o_i,x_i} + \zeta_{i,c_{f_i}}\right) + \xi_{f_i,o_i,x_i} + \epsilon_{i,f_i}$ is the non-pecuniary amenity i receives from her initial firm, and v'_i is the non-pecuniary amenity i receives from her counterfactual firm. We further decompose WTP_i into a term representing the pure effect of earnings changes, and a term representing changes to the amenity value of employment:

$$WTP_i = \underbrace{w'_{o,f'_i} - w_{o,f_i}}_{\text{earnings}} + \underbrace{w_{o,f_i}\left(1 - \exp\left(\frac{v_i - v'_i}{\tau}\right)\right)}_{\text{amenity value of employment}}. \tag{25}$$

Because we assume a competitive, perfectly-elastic output market, the overall welfare effect equals the sum of workers' WTP_i and firms' profit changes.

[38] https://web.archive.org/web/20240523053252/https://www.mbie.govt.nz/assets/summary-of-submissions-consultation-on-employer-assisted-work-visas-and-regional-workforce-planning.pdf.

[39] https://web.archive.org/web/20240523231921/https://www.mbie.govt.nz/assets/c2a0e57f7e/cabinet-paper-report-back-on-consultation-on-proposed-changes-to-immigration-settings-for-international-students.pdf

The result is depicted Figure 10. The overall welfare loss of the Essential Skills visa restrictions is \$410.8m. This is large, equal to about 42% of migrants' baseline earnings. Strikingly, the overall welfare effect is driven by non-pecuniary factors: although expanding migrants' job options would increase aggregate earnings for both migrants and residents, it would also decrease profits, and the net effect on output would be relatively small. The large effect on migrants' non-pecuniary value for their firms is consistent with the result in Section 5 that new residents often switch employers, despite doing so yielding no average benefit to their earnings.

8.3 How the Essential Skills visa system mediates the wage effects of migration

In Section 7, we found that firms' production functions exhibit decreasing returns to scale. As such, migration can decrease residents' wages by decreasing the marginal product of labor. In this subsection, we ask whether the Essential Skills visa system protects residents from the wage effects of increased migration.

In principle, the Essential Skills visa system could either ameliorate or exacerbate the wage effects of migration. When migrants are segregated into certain firms, residents in other firms will be insulated from increased migration. However, increasing migration will have a particularly negative effect on residents in those firms in which migrants can work — both because such firms will suffer from a decreasing marginal product and because, as the migrant share of the firms' employees grows, firms' monopsony power over their employees will grow as well (Amior & Stuhler, 2024). Here, we use our structural model to quantify how the Essential Skills visa system mediates the wage effects of migration.

We focus on a subset of occupations which were not included on an Essential Skills in Demand list during 2018 and which had at least 50 Essential Skills migrants. For each such occupation, we calculate three counterfactual equilibria, in addition to the initial equilibrium which we have already calculated. In counterfactual (1), the number of migrants is increased by 5%. In counterfactual (2), migrants' job options are unrestricted. In counterfactual (3), the number of migrants is increased by 5% *and* migrants' job options are unrestricted.

Our object of interest is the *migration elasticity*: the change in residents' mean log wages, scaled by the change in the log number of workers in the occupation:

$$\frac{\text{change in residents' mean log wage, given a 5\% increase in the number of migrants}}{\text{change in the log number of workers, given a 5\% increase in the number of migrants}}.$$

For each occupation, we calculate two migration elasticities. The *restricted migration elasticity* compares the two equilibria in which migrants job options are restricted (i.e. counterfactual (1) *vs.* the initial equilibrium). The *unrestricted migration elasticity* compares the two equilibria in which migrants job options are unrestricted (i.e. counterfactual (3) *vs.* counterfactual (2)).

These elasticities are compared in Figure 11. Restricted migration elasticities are negative: the Essential Skills visa system does not prevent increased migration from reducing residents' wages. Moreover, in most occupations, the restricted migration elasticity is greater (in absolute value) than the unrestricted migration elasticity: not only does the Essential Skills visa system not protect residents from the wage effects of increased migration, it actually accentuates those effects. In some occupations this difference is large. For example, among bus drivers, we find a restricted migration of -0.27: a 10% increase in the number of bus drivers due to increased migration would decrease the wage of resident bus drivers by 2.7%. Were migrant bus drivers' job choices unrestricted, such an increase in the number of bus drivers would only decrease residents' wages by 1%.

9 Conclusion

How wages are set is a fundamental question that any model of the labor market must address. When wages are bargained individually, the allocation of workers to firms is constrained efficient; the balance of market power matters only for the distribution of surplus (Silbert & Townsend, 2024). Moreover, models of individual wage bargaining typically assume that each worker's wage will be bargained independently of the wage bargained by another — if my employer gives me a raise, its bargain with its other workers will be unaffected.[40]

This paper has argued that assuming individual wage bargaining can be misleading. Our evidence is consistent with a very different wage-setting protocol. Specifically, the firms we study appear to engage in wage-posting: they demonstrate either no ability or no willingness to tailor wages to individual workers. This wage-setting protocol has important consequences for the design of immigration policy: it implies that restrictions on migrants' job options can hurt not only migrants but also the existing residents with whom they work.

More generally, wage-posting links the bargaining power of different groups of workers. As another example, in ongoing work (joint with Savannah Noray) we are studying whether whether the wage gap between men in majority-female occupations and men in majority-male occupations can be explained by womens' weaker job mobility.

This paper has also argued that workers' non-pecuniary preferences are an important policy

[40]There are exceptions. In the Stole and Zwiebel (1996) model, employers value marginal workers in part because marginal workers reduce the bargaining power of inframarginal workers. In the Cullen and Pakzad-Hurson (2023) model, employers are wary of paying one worker a more generous wage because doing so would risk disclosing their ability to pay all workers more.

consideration. When migrants transition from an Essential Skills visa to a resident visa, they very often switch firms but do not, on average, earn any more. By revealed preference, non-pecuniary preferences must be important. When we considered a counterfactual equilibrium in which migrants' job options are unrestricted, we found that these non-pecuniary considerations accounted for 74% of the increase in worker welfare.

References

Ahrens, A., Beerli, A., Hangartner, D., Kurer, S., & Siegenthaler, M. (2023). The labor market effects of restricting refugees' employment opportunities. *IZA Discussion Paper 15901*.

Albert, C. (2021). The labor market impact of immigration: Job creation versus job competition. *American Economic Journal: Macroeconomics*, *13*(1), 35–78.

Allan, C., & Maré, D. C. (2021). Do workers share in firm success? Pass-through estimates for New Zealand. *Motu Working Paper 21-15*.

Amior, M., & Manning, A. (2021). Monopsony and the wage effects of migration. *CEP Discussion Paper 1690*.

Amior, M., & Stuhler, J. (2024). Immigration, monopsony and the distribution of firm pay. *IZA Discussion Paper 16692*.

Azar, J., Berry, S., & Marinescu, I. E. (2022). Estimating labor market power. *NBER Working Paper 30365*.

Bagger, J., Fontaine, F., Postel-Vinay, F., & Robin, J.-M. (2014). Tenure, experience, human capital, and wages: A tractable equilibrium search model of wage dynamics. *American Economic Review*, *104*(6), 1551–1596.

Battisti, M., Felbermayr, G., Peri, G., & Poutvaara, P. (2018). Immigration, search and redistribution: A quantitative assessment of native welfare. *Journal of the European Economic Association*, *16*(4), 1137–1188.

Berger, D., Herkenhoff, K., & Mongey, S. (2022). Labor market power. *American Economic Review*, *112*(4), 1147–1193.

Berry, S. (1994). Estimating discrete-choice models of product differentiation. *The RAND Journal of Economics*, *25*(2), 242–262.

Berry, S., Levinsohn, J., & Pakes, A. (1995). Automobile prices in market equilibrium. *Econometrica*, *63*(4), 841–890.

Bloesch, J., Larsen, B., & Taska, B. (2022). Which workers earn more at productive firms? Position specific skills and individual worker hold-up power.

Borjas, G. J. (1995). The economic benefits from immigration. *Journal of Economic Perspectives*, *9*(2), 3–22.

Borjas, G. J. (2003). The labor demand curve is downward sloping: Reexamining the impact of immigration on the labor market. *The Quarterly Journal of Economics*, *118*(4), 1335–1374.

Borjas, G. J. (2013). The analytics of the wage effect of immigration. *IZA Journal of Migration*, *2*(22), 1–25.

Borjas, G. J., & Edo, A. (2023). *Monopsony, efficiency, and the regularization of undocumented immigrants*.

Brenzel, H., Gartner, H., & Schnabel, C. (2014). Wage bargaining or wage posting? Evidence from the employers' side. *Labour Economics*, *29*, 41–48.

Burdett, K., & Mortensen, D. T. (1998). Wage differentials, employer size, and unemployment. *International Economic Review*, *39*(2), 257–273.

Cahuc, P., Postel-Vinay, F., & Robin, J.-M. (2006). Wage bargaining with on-the-job search: Theory and evidence. *Econometrica*, *74*(2), 323–364.

Caldwell, S., & Danieli, O. (2024). Outside options in the labour market. *The Review of Economic Studies*, *91*(6), 3286–3315.

Caldwell, S., Haegele, I., & Heining, J. (2024). Bargaining and inequality in the labor market.

Caldwell, S., & Harmon, N. (2019). Outside options, bargaining, and wages: Evidence from coworker networks.

Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2011). Robust inference with multiway clustering. *Journal of Business & Economic Statistics*, *29*(2), 238–249.

Caplin, A., & Nalebuff, B. (1991). Aggregation and imperfect competition: On the existence of equilibrium. *Econometrica*, *59*(1), 25–59.

Card, D. (1990). The impact of the mariel boatlift on the miami labor market. *Industrial and Labor Relations Review*, *43*(2), 245–257.

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

Card, D., & Hyslop, D. R. (2005). Estimating the effects of a time-limited earnings subsidy for welfare-leavers. *Econometrica*, *73*(6), 1723–1770.

Carrozzo, S. (2022). The labor market effects of an unexpected amnesty for undocumented workers.

Carvalho, M., Galindo da Fonseca, J., & Santarrosa, R. (2023). How are wages determined? A quasi-experimental test of wage determination theories. *Rimini Centre for Economic Analysis Working Paper 2023-08*.

Chan, M., Kroft, K., Mattana, E., & Mourifié, I. (2024). An empirical framework for matching with imperfect competition. *NBER Working Paper 32493*.

Chassamboulli, A., & Palivos, T. (2013). The impact of immigration on the employment and wages of native workers. *Journal of Macroeconomics*, *38*, 19–34.

Chassamboulli, A., & Palivos, T. (2014). A search-equilibrium approach to the effects of immigration on labor market outcomes. *International Economic Review*, *55*(1), 111–129.

Chassamboulli, A., & Peri, G. (2015). The labor market effects of reducing the number of illegal immigrants. *Review of Economic Dynamics*, *18*(4), 792–821.

Chib, S. (2001). Markov chain Monte Carlo methods: Computation and inference. *Handbook of Econometrics*, *5*, 3569–3649.

Cullen, Z. B., & Pakzad-Hurson, B. (2023). Equilibrium effects of pay transparency. *Econometrica*, *91*(3), 765–802.

De Chaisemartin, C., & d'Haultfoeuille, X. (2024). Difference-in-differences estimators of intertemporal treatment effects. *Review of Economics and Statistics*, 1–45.

De Chaisemartin, C., & d'Haultfoeuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, *110*(9), 2964–2996.

Delabastita, V., & Rubens, M. (2022). Colluding against workers.

Di Porto, E., Martino, E. M., & Naticchioni, P. (2018). Back to black? The impact of regularizing migrant workers. *CSEF Working Papers 517*.

Doran, K., Gelber, A., & Isen, A. (2022). The effects of high-skilled immigration policy on firms: Evidence from visa lotteries. *Journal of Political Economy*, *130*(10), 2501–2533.

Dube, A., Jacobs, J., Naidu, S., & Suri, S. (2020). Monopsony in online labor markets. *American Economic Review: Insights*, *2*(1), 33–46.

Dustmann, C., Schönberg, U., & Stuhler, J. (2017). Labor supply shocks, native wages, and the adjustment of local employment. *The Quarterly Journal of Economics*, *132*(1), 435–483.

D'Haultfoeuille, X. (2011). On the completeness condition in nonparametric instrumental problems. *Econometric Theory*, *27*(3), 460–471.

Elias, F., Monras, J., & Vázquez-Grenno, J. (2025). Understanding the effects of granting work permits to undocumented immigrants. *Journal of Labor Economics*, *43*(3), 763–802.

Fabling, R., & Maré, D. C. (2019). Improved productivity measurement in New Zealand's longitudinal business database. *Motu Working Paper 19-03*.

Freyaldenhoven, S., Hansen, C., Pérez, J. P., & Shapiro, J. M. (2021). Visualization, identification, and estimation in the linear panel event-study design. *NBER Working Paper 29170*.

Freyaldenhoven, S., Hansen, C., & Shapiro, J. M. (2019). Pre-event trends in the panel event-study design. *American Economic Review*, *109*(9), 3307–3338.

Gibson, J., McKenzie, D., Rohorua, H., & Stillman, S. (2019). The long-term impact of international migration on economic decision-making: Evidence from a migration lottery and lab-in-the-field experiments. *Journal of Development Economics*, *138*, 99–115.

Groetsch, C. W. (2007). Integral equations of the first kind, inverse problems and regularization: A crash course. In *Journal of physics: Conference series* (Vol. 73).

Hall, R. E., & Krueger, A. B. (2012). Evidence on the incidence of wage posting, wage bargaining, and on-the-job search. *American Economic Journal: Macroeconomics*, *4*(4), 56–67.

Heckman, J. J., Smith, J., & Clements, N. (1997). Making the most out of programme evaluations and social experiments: Accounting for heterogeneity in programme impacts. *The Review of Economic Studies*, *64*(4), 487–535.

Kossoudji, S. A., & Cobb-Clark, D. A. (2002). Coming out of the shadows: Learning about legal status and wages from the legalized population. *Journal of Labor Economics*, *20*(3), 598–628.

Kroft, K., Luo, Y., Mogstad, M., & Setzler, B. (2020). Imperfect competition and rents in labor and product markets: The case of the construction industry. *NBER Working Paper 27325*.

Lachowska, M., Mas, A., Saggio, R., & Woodbury, S. A. (2022). Wage posting or wage bargaining? a test using dual jobholders. *Journal of Labor Economics*, *40*(S1), S469–S493.

Lamadon, T., Mogstad, M., & Setzler, B. (2022). Imperfect competition, compensating differentials, and rent sharing in the US labor market. *American Economic Review*, *112*(1), 169–212.

Lee, D. S. (2009). Training, wages, and sample selection: Estimating sharp bounds on treatment effects. *Review of Economic Studies*, *76*(3), 1071—1102.

MacKay, A., & Miller, N. H. (2025). Estimating models of supply and demand: Instruments and covariance restrictions. *American Economic Journal: Microeconomics*, *17*(1), 238—81.

Manning, A. (2005). *Monopsony in motion: Imperfect competition in labor markets*. Princeton University Press.

Maré, D. C., & Hyslop, D. (2021). Minimum wages in New Zealand: Policy and practice in the 21st century. *IZA Discussion Paper 14302*.

McCulloch, R., & Rossi, P. E. (1994). An exact likelihood analysis of the multinomial probit model. *Journal of Econometrics*, *64*(1–2), 207–240.

McKenzie, D., Stillman, S., & Gibson, J. (2010). How important is selection? Experimental vs. non-experimental measures of the income gains from migration. *Journal of the European Economic Association*, *8*(4), 913–945.

Milgrom, P., & Shannon, C. (1994). Monotone comparative statics. *Econometrica*, *62*(1), 157–180.

Monras, J. (2020). Immigration and wage dynamics: Evidence from the Mexican peso crisis. *Journal of Political Economy*, *128*(8), 3017–3089.

Naidu, S., Nyarko, Y., & Wang, S.-Y. (2016). Monopsony power in migrant labor markets: Evidence from the United Arab Emirates. *Journal of Political Economy*, *124*(6), 1735–1792.

Nash, J. F. (1950). The bargaining problem. *Econometrica*, *18*(1), 155–162.

New Zealand Productivity Commission. (2022). *Immigration by the numbers* (Tech. Rep.).

OECD. (2021). *OECD/AIAS database on institutional characteristics of trade unions, wage setting, state intervention and social pacts.* `https://stats.oecd.org/Index.aspx?DataSetCode=TUD`. (Accessed: 2024-04-18.)

OECD. (2024). *Minimum relative to average wages of full-time workers.* `https://stats.oecd.org/viewhtml.aspx?datasetcode=MIN2AVE`. (Accessed: 2024-04-18.)

Pissarides, C. A. (2000). *Equilibrium unemployment theory*. MIT Press.

Roussille, N., & Scuderi, B. (2023). Bidding for talent: A test of conduct in a high-wage labor market. *IZA Discussion Paper 16352*.

Rubinstein, A. (1982). Perfect equilibrium in a bargaining model. *Econometrica*, *50*(1), 97–109.

Silbert, J., & Townsend, W. (2024). Job matching without price discrimination.

Sokolova, A., & Sorensen, T. (2021). Monopsony in labor markets: A meta-analysis. *ILR Review*, *74*(1), 27–55.

Stillman, S., McKenzie, D., & Gibson, J. (2009). Migration and mental health: Evidence from a natural experiment. *Journal of Health Economics*, *28*(3), 677–687.

Stole, L. A., & Zwiebel, J. (1996). Intra-firm bargaining under non-binding contracts. *The Review of Economic Studies*, *63*(3), 375–410.

Taber, C., & Vejlin, R. (2020). Estimation of a Roy/search/compensating differential model of the labor market. *Econometrica*, *88*(3), 1031–1069.

Townsend, W. (in press). A novel algorithm for drawing nested extreme value random variables. *Journal of Choice Modelling*.

Vangel, M. G. (1992). *Iterative algorithms for integral equations of the first kind with applications to statistics*. Harvard University.

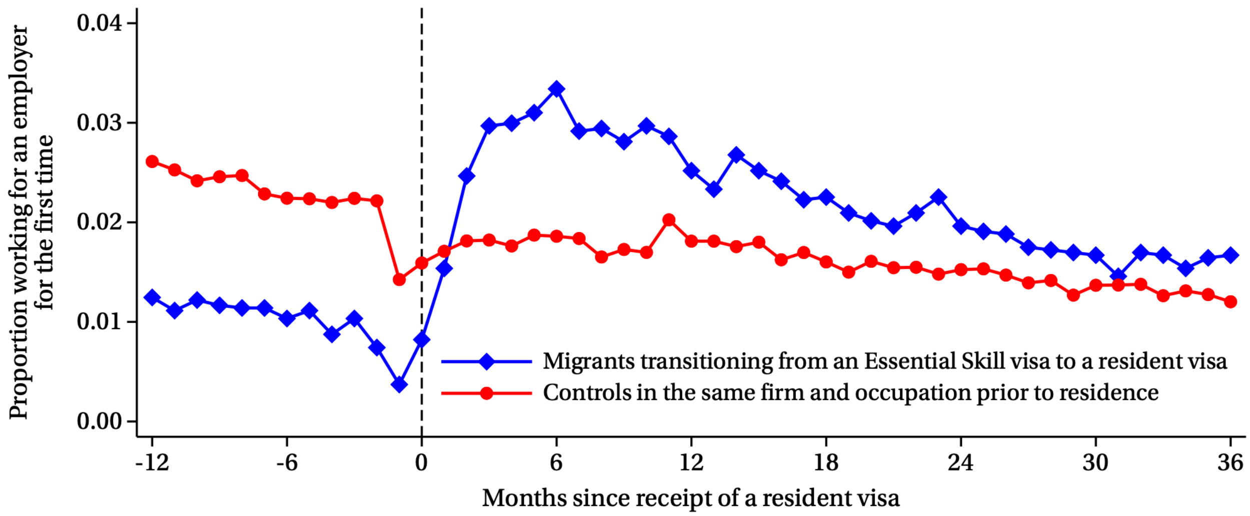
Wang, X. (2021). US permanent residency, job mobility, and earnings. *Journal of Labor Economics*, *39*(3), 639–671.

Yeh, C., Macaluso, C., & Hershbein, B. (2022). Monopsony in the US labor market. *American Economic Review*, *112*(7), 2099–2138.

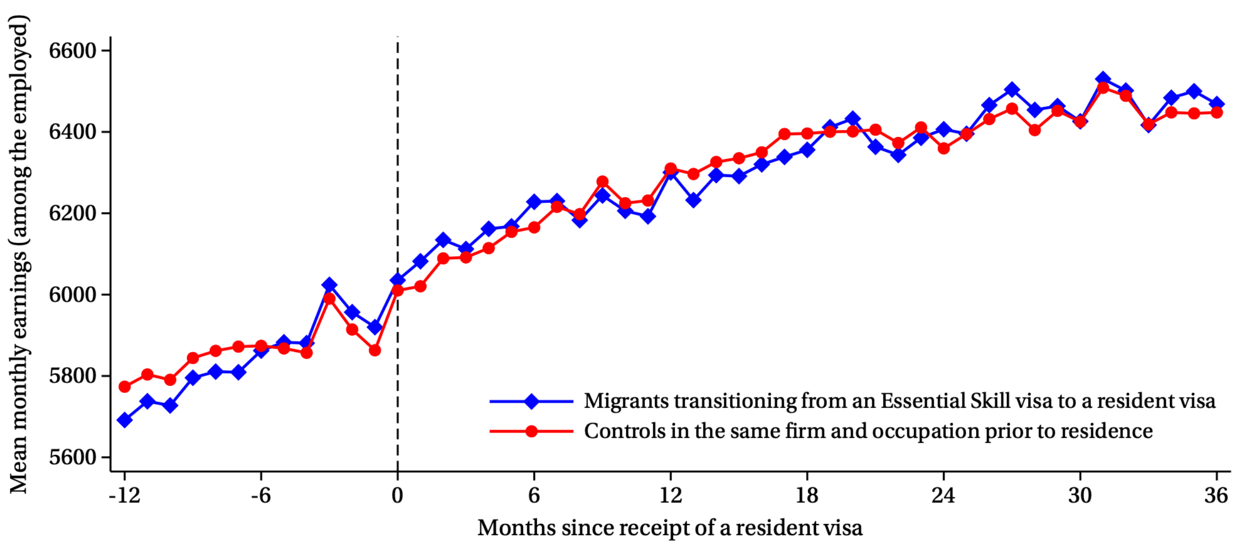
Figure 1: The relationship between a worker's earnings and her colleagues' visa status

This binned scatter plot presents mean employee earnings, conditional on both a worker's own visa status and on the proportion of the workers in the modal worker's occupation × firm cell who hold an Essential Skills visa. Earnings are measured in 2020 New Zealand dollars. The figure is constructed using both the 2013 and 2018 population censuses; the sample are full-time paid employees with a unique employer. See Section 3 for data sources and Section 4 for details on the analytic sample.

Figure 2: How job-switching and earnings change when a migrant becomes a resident



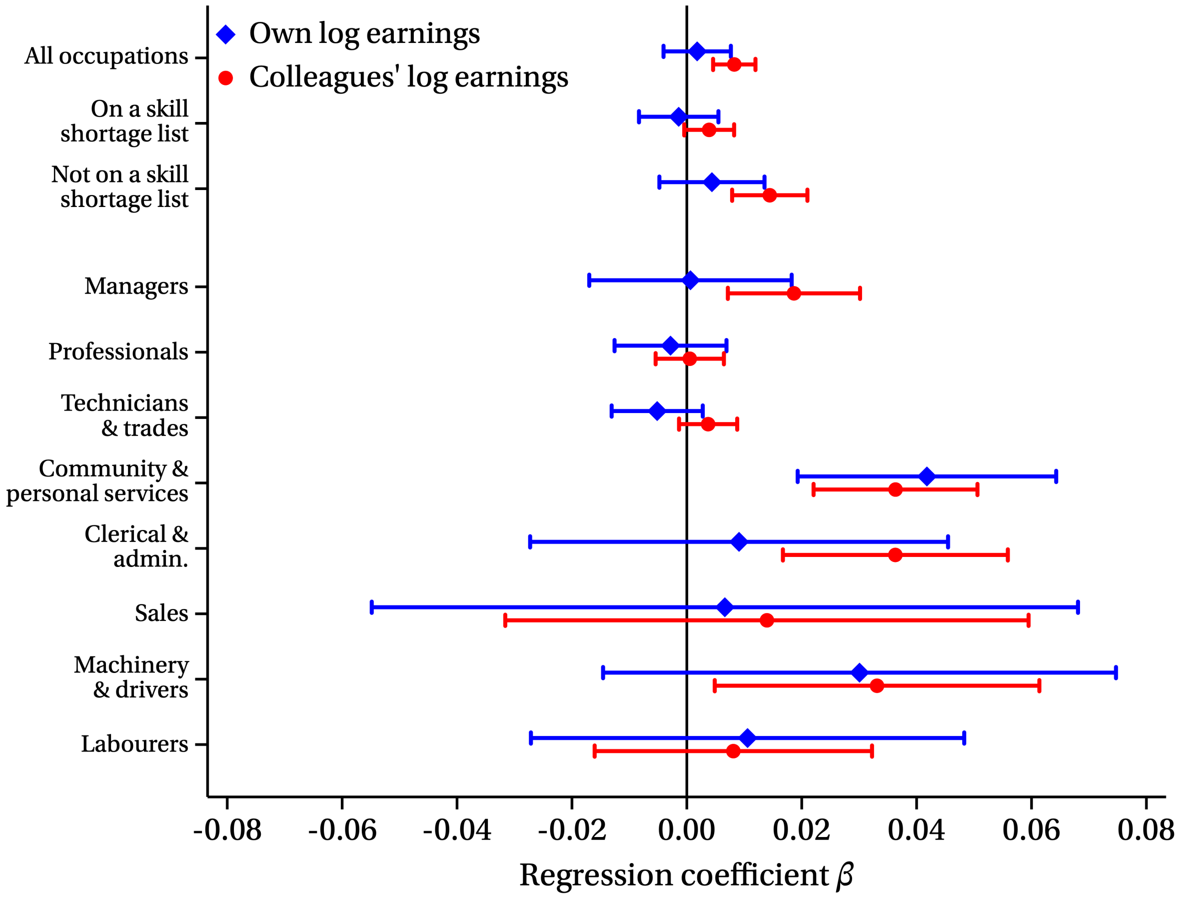
Panel A: Job-switching



Panel B: Earnings

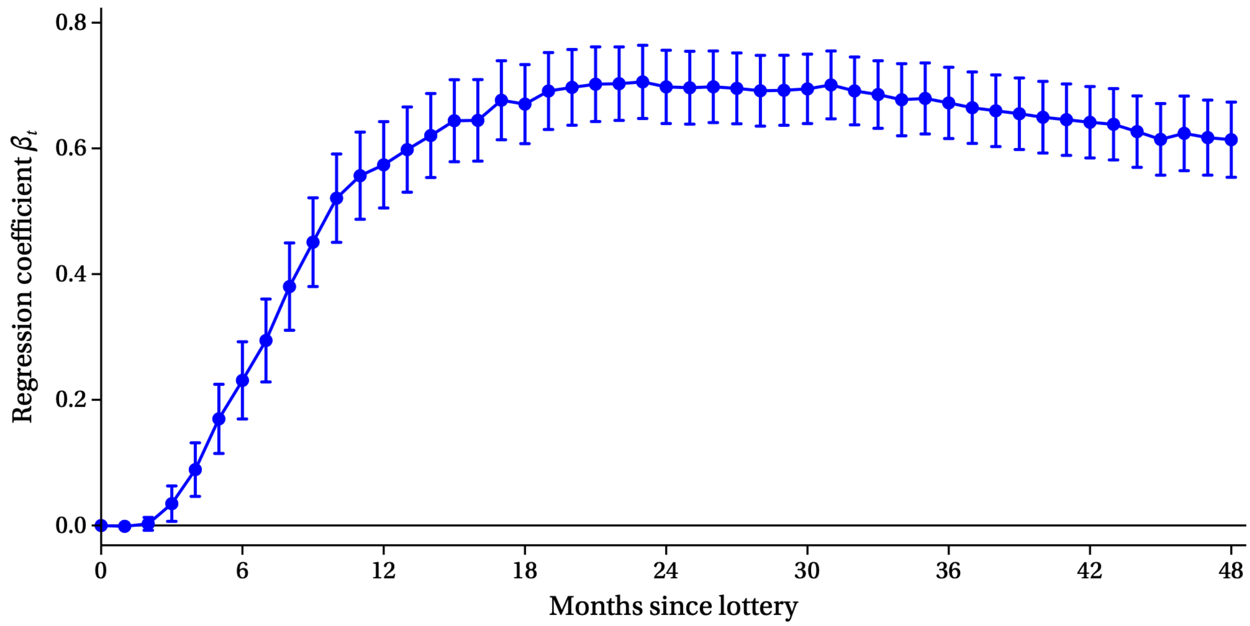
This figure depicts how job-switching (Panel A) and earnings (Panel B) change around the month that a migrant transitions from an Essential Skills visa to a resident visa. The dependent variable in Panel A indicates a worker being employed at a particular firm for the first time. The dependent variable in Panel B is monthly labor earnings (in 2020 New Zealand dollars). Both figures comprise only workers who held an Essential Skills visa for at least 18 months prior to receiving a resident visa; Panel B additionally restricts to workers with positive earnings. The red series comprises workers who worked in the same firm and occupation as the modal migrant in the month prior to the modal migrant receiving a resident visa. The dip in the baseline month in Panel A is due to our requirement that both the new resident and her controls have a unique firm in that month. See Section 3 for data sources and Subsection 5.1 for details on the analytic sample.

Figure 3: How receiving a resident visa affects a worker's earnings

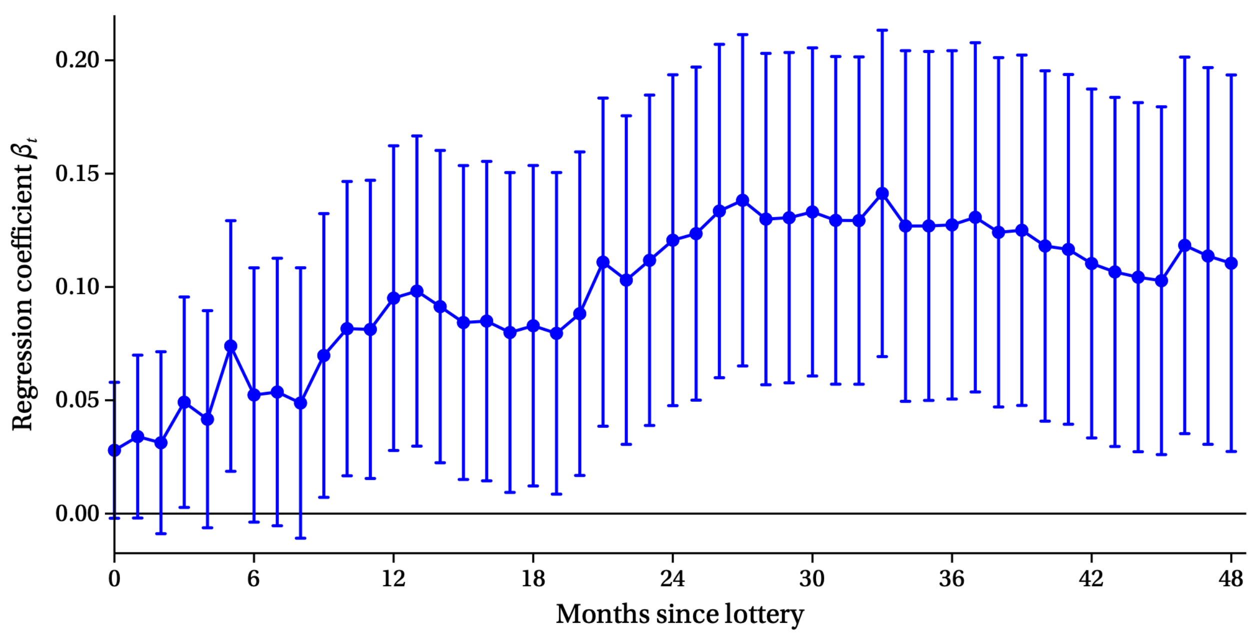


This figure depicts estimated effects of receiving a resident visa on earnings using the matched-control estimator of Equation (2). The dependent variable in the blue diamond series is a worker's monthly log earnings. The dependent variable in the red circle series is the average monthly log earnings among a worker's colleagues. Both series comprise only workers who held an Essential Skills visa for at least 18 months prior to receiving a resident visa and who have positive earnings. The red series additionally restricts to workers with a unique firm and excludes the first and last month of employment at a firm. Bars represent 95% confidence intervals. See Section 3 for data sources and Subsection 5.1 for details on the analytic sample and estimator. The underlying data for this figure is available in Appendix Table A2.

Figure 4: How winning a resident visa lottery affects a worker's visa status, job-switching and earnings

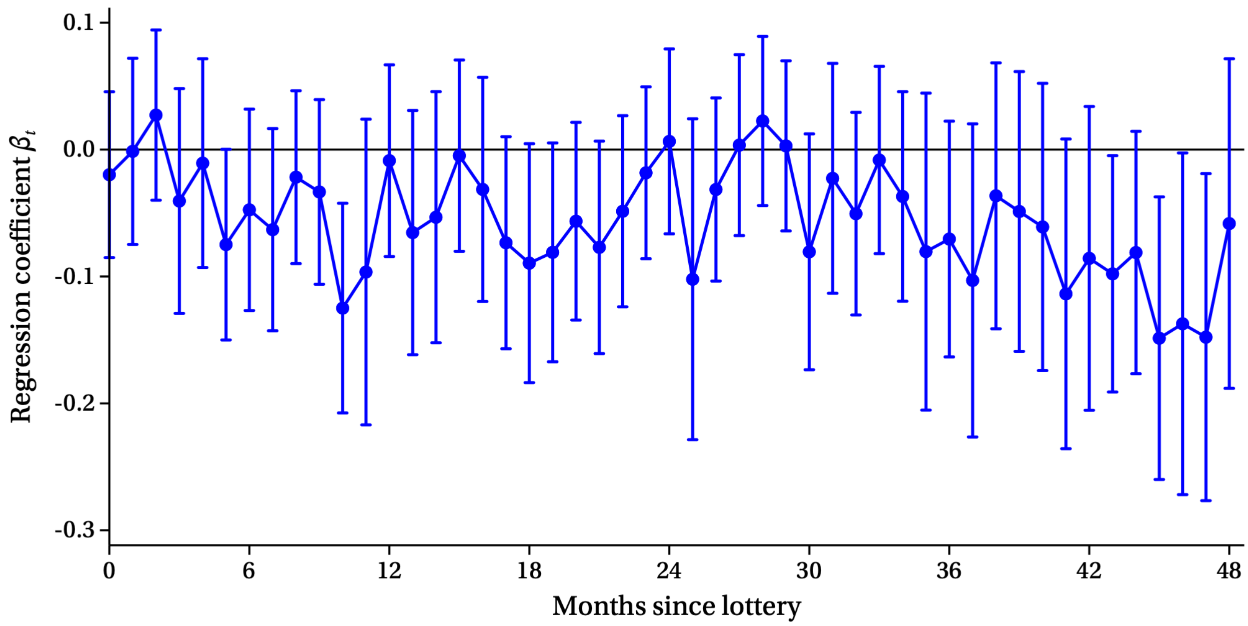


Panel A: Effects on visa status



Panel B: Effects on job-switching

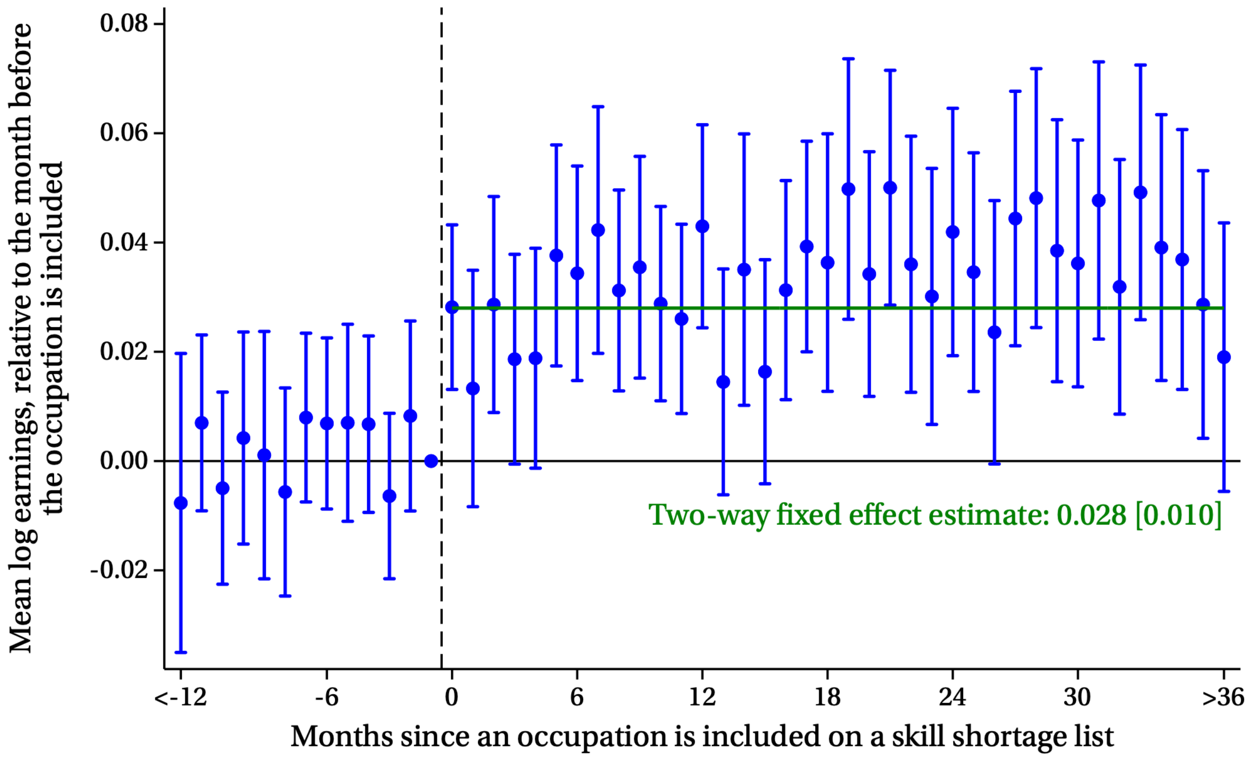
Figure 4 (continued): How winning a resident visa lottery affects a worker's visa status, job-switching and earnings



Panel C: Effects on log earnings

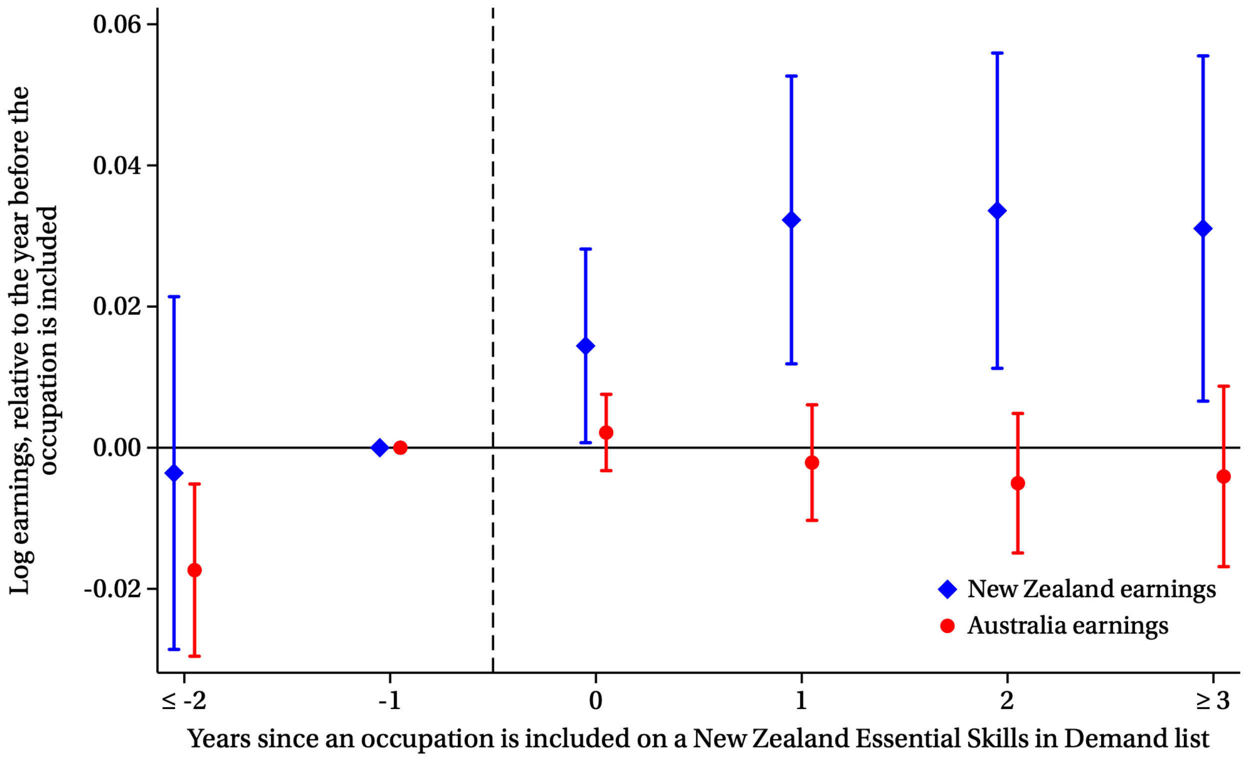
The figure presents regression estimates of Equation (4) to study the effects of winning a resident visa lottery, varying the month t at which outcomes are measured. In Panel A, the dependent variable is an indicator for having obtained a resident visa t months after the lottery. In Panel B, the dependent variable is an indicator for having worked for a new firm at any point during the t months since the lottery. In Panel C, the dependent variable is log earnings t months after the lottery. Bars represent 95% confidence intervals. See Section 3 for data sources and Subsection 5.2 for details on the analytic sample and estimator.

Figure 5: How earnings change when an occupation is included on an Essential Skills in Demand list



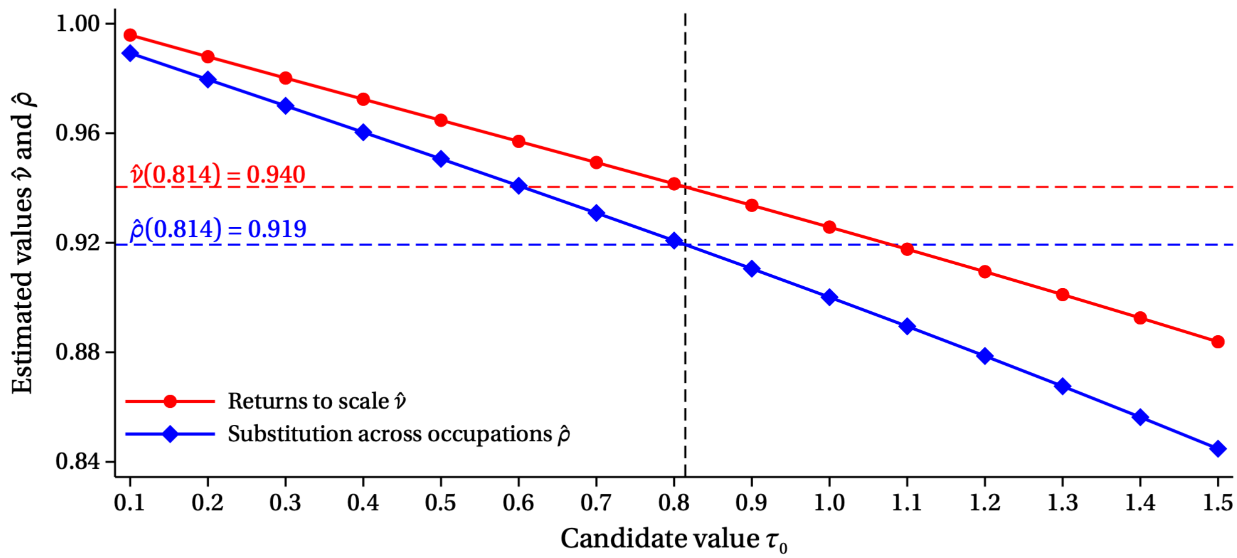
This figure reports coefficient estimates from the Freyaldenhoven et al. (2021) linear panel event-study model, asking how log earnings in an occupation typically change when the occupation is added to or removed from an Essential Skills in Demand list. Bars represent 95% confidence intervals. See Section 3 for data sources and Section 6 for details on the analytic sample.

Figure 6: How earnings in New Zealand and Australia change when an occupation is included on a New Zealand Essential Skills in Demand list

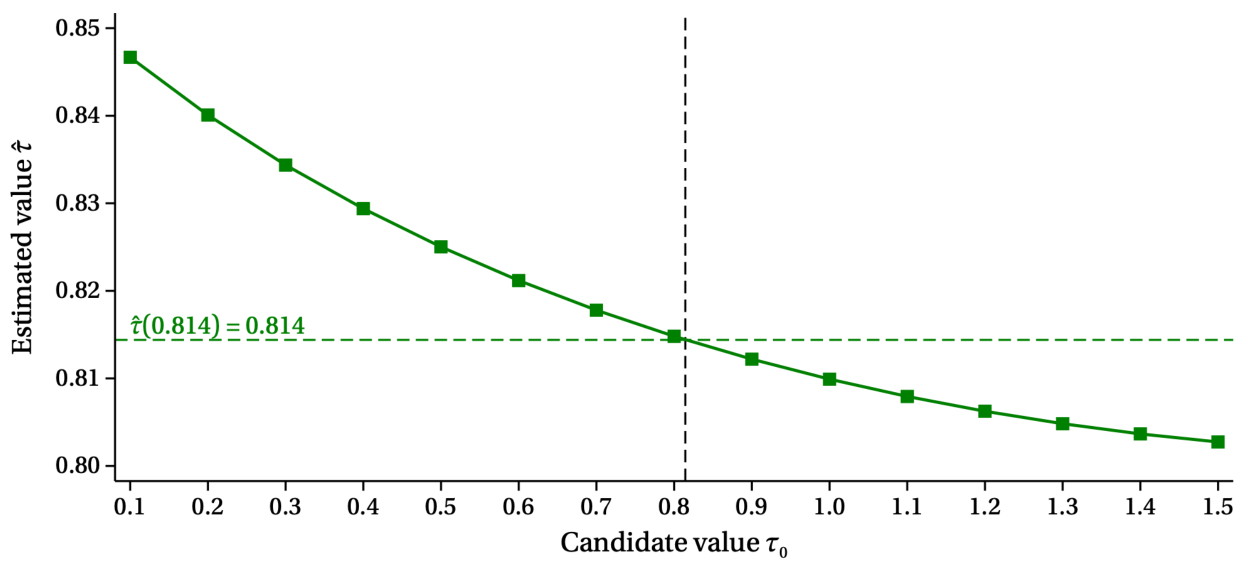


This figure reports coefficient estimates from the Freyaldenhoven et al. (2021) linear panel event-study model, asking how log earnings in an occupation (in New Zealand or in Australia) typically change when the occupation is added to or removed from an New Zealand Essential Skills in Demand list. Bars represent 95% confidence intervals. See sections 3 and 6 for data sources and Section 6 for details on the analytic sample.

Figure 7: The fixed-point estimation procedure



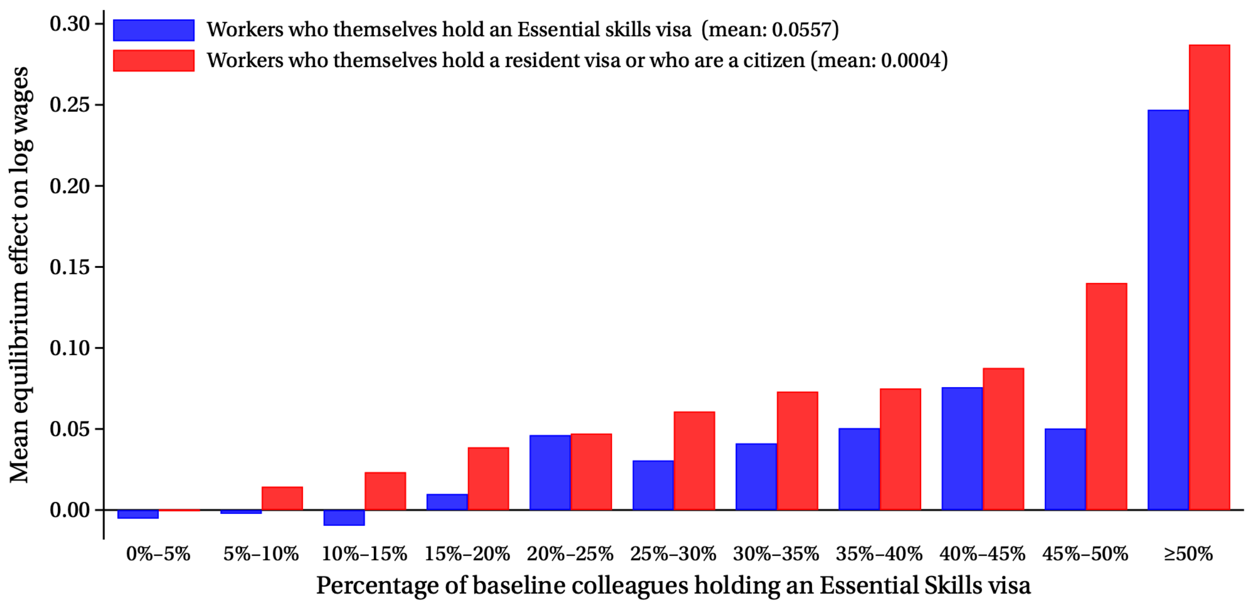
Panel A: Estimates of ν and ρ



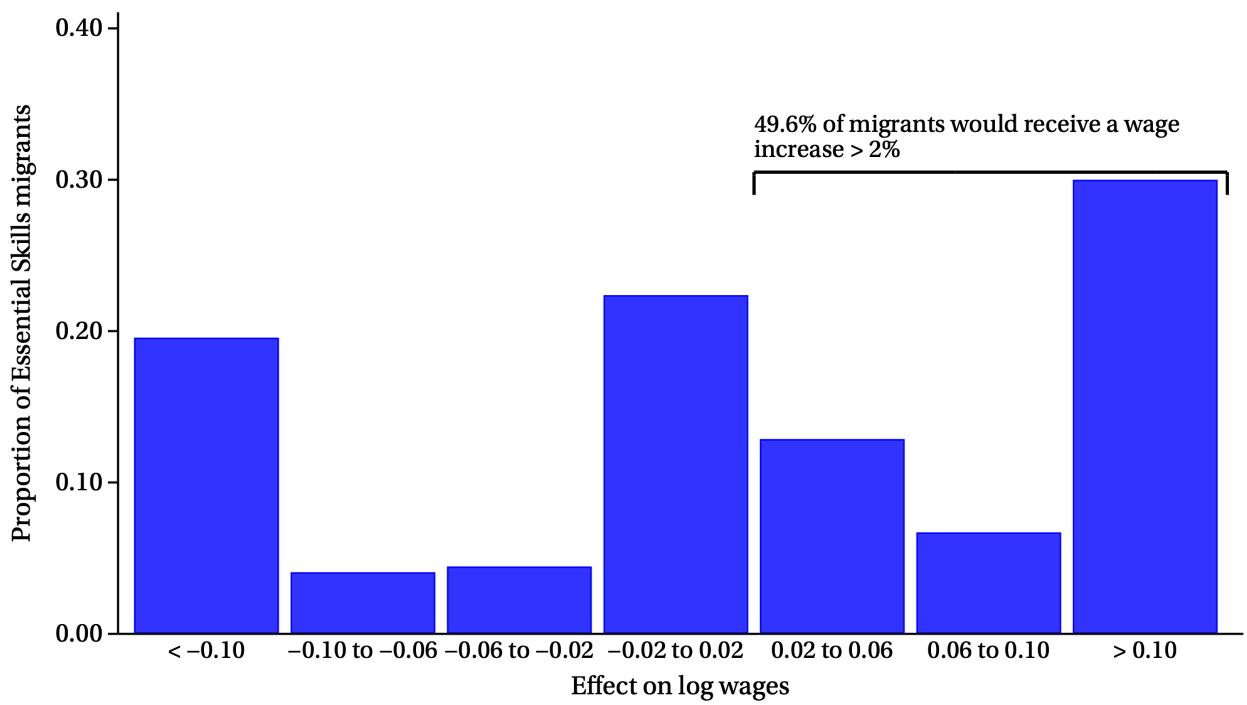
Panel B: Estimates of τ

This figure illustrates the fixed-point estimation procedure presented in Section 7. For each candidate value of the labor supply parameter τ_0, Panel A depicts the implied estimates of the structural production parameters $\hat{\nu}$ and $\hat{\rho}$. In turn, these production parameters imply an estimate of the labor supply parameter $\hat{\tau}$, which are each depicted in Panel B. Our point estimate — the fixed point $\hat{\tau} = \tau_0 = 0.814$ — is depicted in dashed lines. All estimates fix $\lambda = 0.131$.

Figure 8: How expanding migrants' job options affects wages

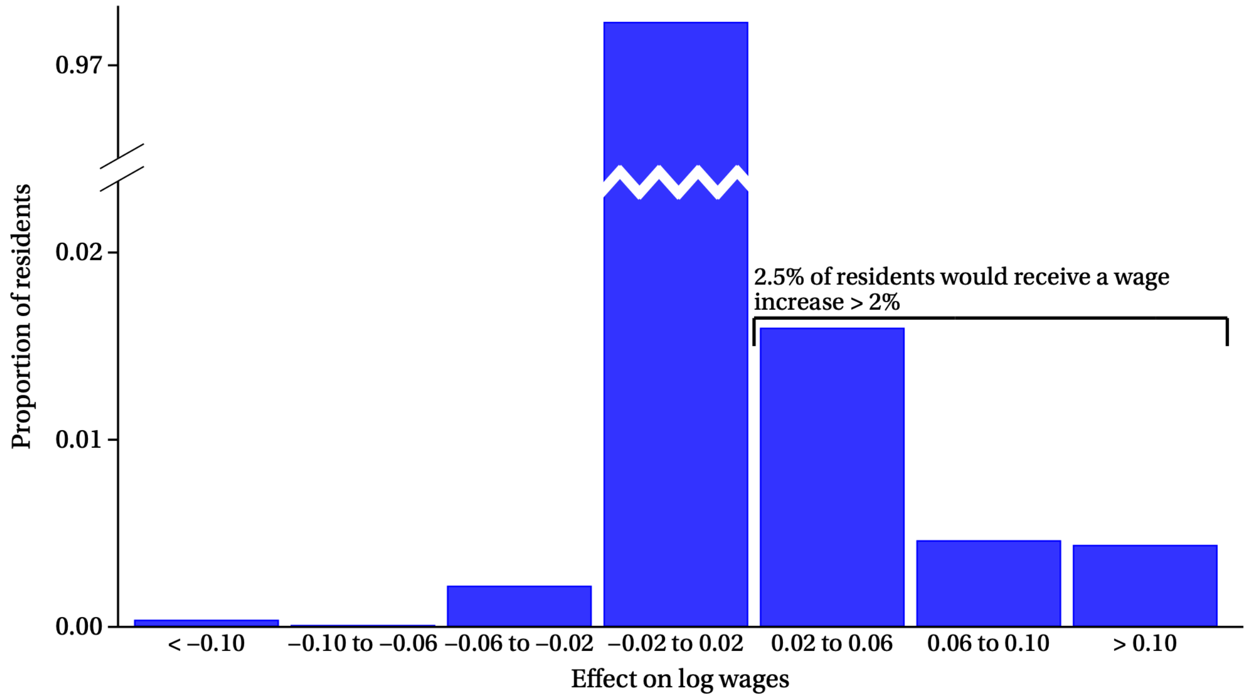


Panel A: Mean effects by own visa status and the visa status colleagues



Panel B: The distribution of effects on migrants' log wages

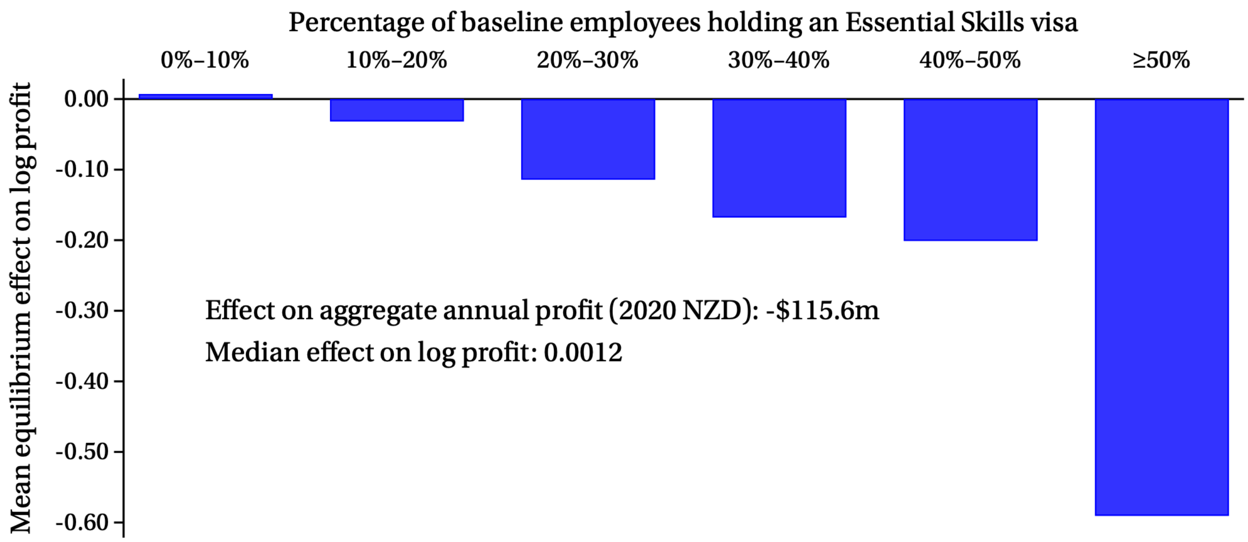
Figure 8 (continued): How expanding migrants' job options affects wages



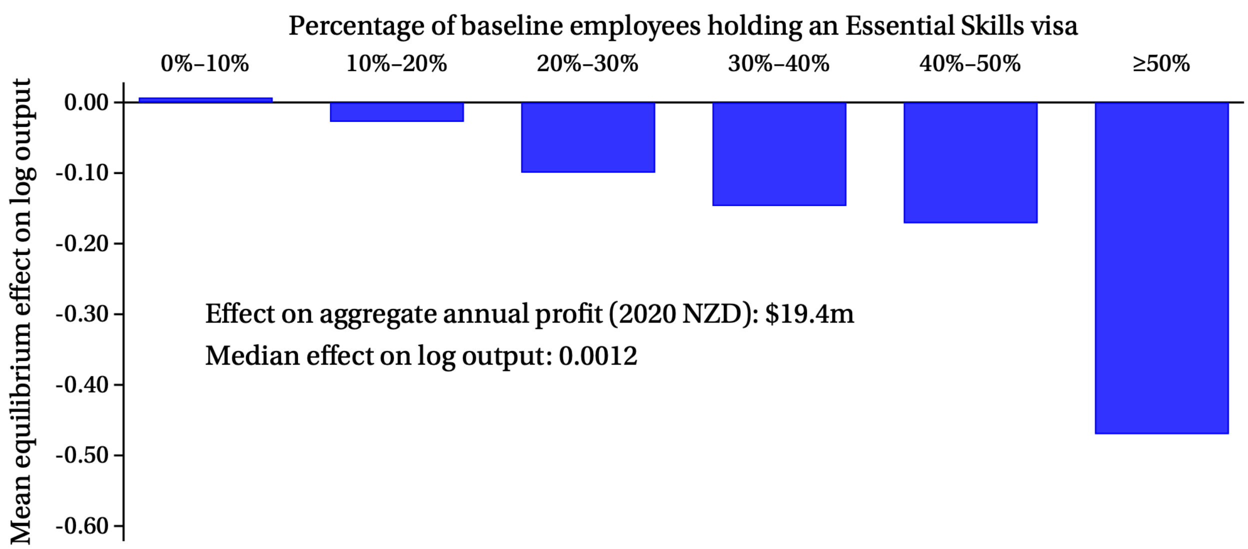
Panel C: The distribution of effects on residents' log wages

This figure depicts how allowing migrants to work at any firm would affect workers' log wages, using the equilibrium simulations discussed in Section 8. Panel A depicts the mean effect on log wages, by a worker's own visa status and by the percentage of their baseline colleagues who themselves hold an Essential Skills visa. Panel B depicts the distribution of effects among Essential Skills migrants while Panel C depicts the distribution of effects among residents. Workers whose occupation is already included on a nation-wide Essential Skills in Demand list are excluded from the population used to generate the figure.

Figure 9: How expanding migrants' job options affects firm profit and output



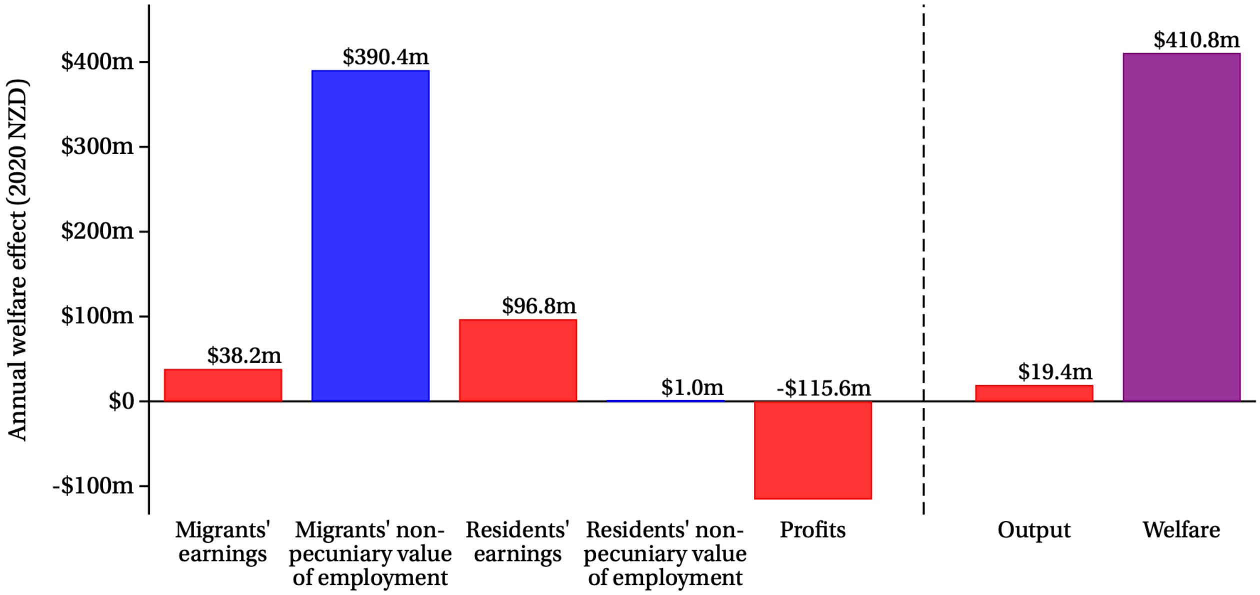
Panel A: Mean effects on log profit, by visa status among employees



Panel B: Mean effects on log output, by visa status among employees

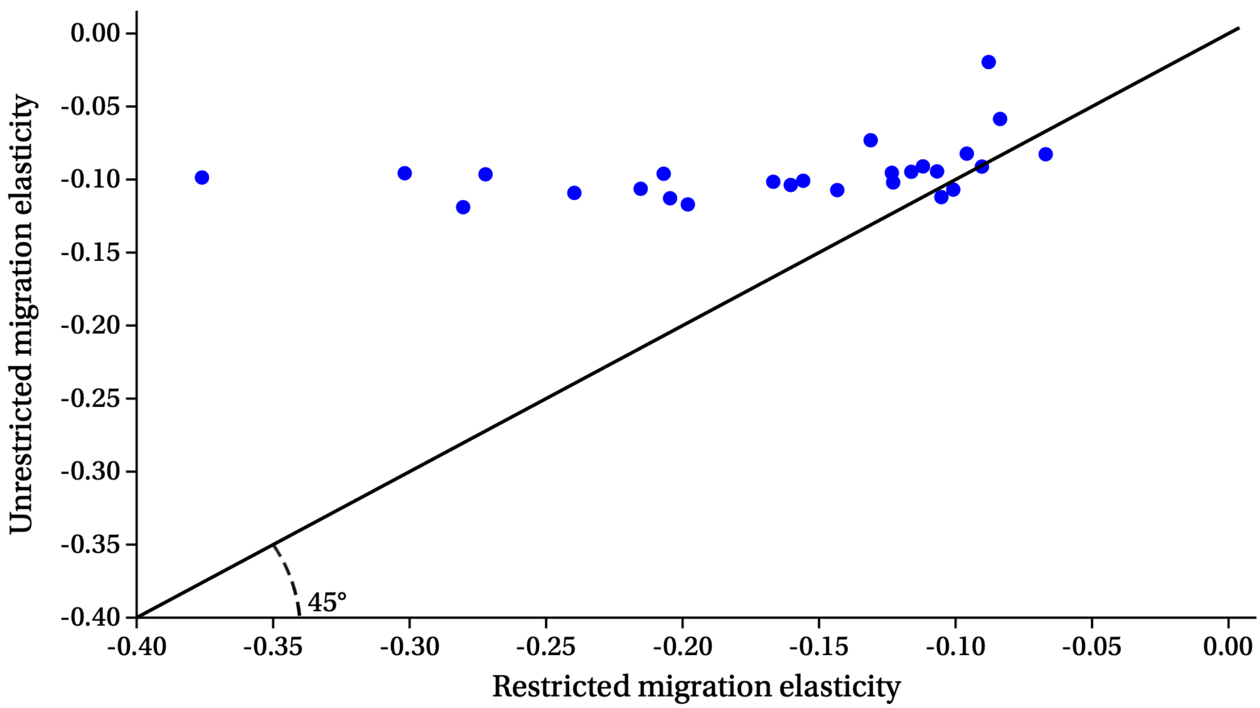
This figure depicts how allowing migrants to work at any firm would affect firms' profits and output, using the equilibrium simulations discussed in Section 8. Panel A depicts the mean effect on log profit by the percentage of the firm's employees who hold an Essential Skills visa. Panel B depicts the mean effect on log output by the percentage of the firm's employees who hold an Essential Skills visa.

Figure 10: How expanding migrants' job options affects aggregate welfare



This figure depicts how allowing migrants to work at any firm would affect aggregate welfare and its subcomponents, using the equilibrium simulations discussed in Section 8. Worker welfare is measured as in equations (24) and (25).

Figure 11: How the Essential Skills visa system mediates the effects of migration



This figure depicts how increasing the number of migrants would affect residents' wages — and how this effect would differ, were migrants' job options unrestricted. Both axes depict a migration elasticity: the effect of increasing the number of workers in an occupation by 1%, through increased immigration, on residents' mean wages. When simulating the restricted migration elasticity, reported on the x-axis, we allow migrants to only work at certain firms. When simulating the unrestricted migration elasticity, reported on the y-axis, we allow migrants to work at any firm. These simulations are described more fully in Subsection 8.3.

Table 1: Summary statistics

Panel A: All worker-month observations				
	By visa status:			All workers
	Essential Skills	Other temporary visa	Resident or citizen	
Proportion male	0.715	0.561	0.496	0.502
Mean real earnings	5067.0	3564.2	5215.4	5135.8
Proportion living in:				
Auckland	0.315	0.373	0.266	0.271
any major urban area	0.540	0.578	0.510	0.514
Number of worker-month observations	4069371	13074444	261511728	278655552
Number of distinct workers	205875	872676	3386520	3840834

Panel B: During the 2018 Population Census				
	By visa status:			All workers
	Essential Skills	Other temporary visa	Resident or citizen	
Proportion male	0.726	0.522	0.489	0.494
Mean age	35.1	31.0	43.1	42.5
Proportion living in:				
Auckland	0.410	0.528	0.355	0.362
any major urban area	0.842	0.918	0.883	0.883
Proportion born in:				
New Zealand	0.001	0.003	0.716	0.679
United Kingdom	0.105	0.092	0.065	0.066
India	0.124	0.205	0.025	0.033
China	0.061	0.131	0.021	0.026
South Africa	0.058	0.038	0.022	0.023
Fiji	0.051	0.026	0.017	0.018
Australia	—	0.005	0.018	0.017
Philippines	0.266	0.093	0.018	0.024
Proportion by highest qualification:				
none	0.057	0.043	0.116	0.113
high school	0.386	0.287	0.372	0.369
trade qualification	0.177	0.147	0.196	0.194
bachelor's degree	0.242	0.291	0.180	0.185
postgraduate degree	0.138	0.232	0.135	0.139
Proportion paid employees	0.994	0.989	0.949	0.950
Proportion working full-time in their primary job	0.958	0.750	0.801	0.801
Proportion with a secondary job	0.015	0.056	0.069	0.068
Mean real labor income in prior year	50794.2	31668.4	61337.1	60020.0
Mean num. months of income in prior year	10.4	8.4	11.2	11.0
Mean num. employers in prior year	1.3	1.7	1.4	1.4
Number of workers	19152	56355	1435512	1511019
Prop. of sample matched to a Census form	0.759	0.658	0.926	0.909

Table 1 describes our sample of worker-month observations. Panel A includes all person-months with positive earnings, while Panel B includes the subset of such observations from March 2018 (when the 2018 Population Census was conducted) who filled out a 2018 Census form. Visa status is from administrative records. In Panel A, gender and location are from both survey and administrative records, while earnings are from administrative tax records. In Panel B, all variables are from the 2018 Population Census, except for labor income and employer counts, which are from administrative tax records. Earnings are expressed in 2020 New Zealand dollars.

Table 2: The most common occupations among Essential Skills migrants

Occupation	Num. Essential Skills migrants	Prop. of Essential Skills migrants in this occupation	Prop. of occupation who are an Essential Skills migrant	Prop. of occupation who are an Essential Skills migrant, among Essential Skills migrants' firms
Dairy Cattle Farm Worker	951	0.056	0.341	0.856
Chef	918	0.054	0.095	0.634
Carpenter	903	0.053	0.166	0.707
Aged or Disabled Carer	633	0.037	0.179	0.506
Dairy Cattle Farmer	525	0.031	0.090	0.701
Cafe or Restaurant Manager	438	0.026	0.106	0.789
Retail Supervisor	420	0.025	0.126	0.825
Retail Manager (General)	366	0.022	0.024	0.704
Metal Fabricator	261	0.015	0.113	0.712
Personal Care Assistant	255	0.015	0.026	0.204
Cook	252	0.015	0.118	0.911
Registered Nurse (Aged Care)	180	0.011	0.067	0.308
Truck Driver (General)	171	0.010	0.010	0.396
Commercial Housekeeper	165	0.010	0.129	0.669
Motor Mechanic (General)	162	0.010	0.022	0.475
Sales Assistant (General)	159	0.009	0.005	0.548
Resident Medical Officer	150	0.009	0.041	0.080
Diesel Motor Mechanic	147	0.009	0.043	0.545
Massage Therapist	141	0.008	0.427	0.871
Waiter	129	0.008	0.051	0.595
All occupations	16992	1.000	0.017	0.635

Table 2 lists the 20 most common occupations for Essential Skills migrants, during March 2018. Counts and proportions are calculated using all workers with positive earnings who can be assigned an occupation either because they completed a census form or because their occupation is included in administrative immigration records. To construct the final column, we first calculate the proportion of each occupation-firm cell who hold an Essential Skills visa, and then take the average of these proportions weighting each cell by its number of Essential Skills migrants.

Table 3: Cross-sectional wage regressions

	(1)	(2)	(3)	(4)	(5)	(6)
Sample: 2013 Population Census						
Slope coefficient on Essential Skills visa	-0.194	-0.184	-0.048	-0.025	-0.010	-0.032
	[0.006]	[0.005]	[0.004]	[0.005]	[0.006]	[0.006]
Number of workers	714915	696870	679533	696870	679533	696648
Sample: 2018 Population Census						
Slope coefficient on Essential Skills visa	-0.222	-0.189	-0.049	-0.048	-0.004	-0.026
	[0.003]	[0.003]	[0.003]	[0.004]	[0.005]	[0.005]
Number of workers	772842	770367	750789	770367	750789	753039
Controls:	No	Yes	Yes	Yes	Yes	No
Fixed effects:	None	None	Occupation	Firm	Occ. × firm	Occ. × firm

Table 3 reports cross-sectional regressions in which the dependent variable is an employee's log earnings (in the 12 months prior to either the 2013 Population Census or the 2018 Population Census). These regressions are estimated using full-time paid employees with a unique employer (see Section 4 for a detailed description of the sample). The controls included in columns (2)-(5) are log hours of work, indicators for decadal age bins, and indicators for the worker having as their highest qualification a high school certificate, a trade certificate or diploma, a bachelor's degree, or a postgraduate degree. (Estimated slope coefficients for these controls, along with R^2 values, are reported in Appendix Table A1). Standard errors (robust to heteroskedasticity) are in brackets.

Table 4: Resident visa lottery placebo analysis

	Control mean	Slope coefficient	Standard error	P. value	Num. workers	Num. obs.
Age	34.805	-0.436	[0.553]	0.430	3465	6123
Number of dependants	1.894	0.021	[0.154]	0.890	2985	5070
Months in New Zealand	44.441	3.759	[2.485]	0.130	3465	6123
Male	0.793	-0.008	[0.032]	0.800	3456	6108
In Auckland	0.456	-0.019	[0.040]	0.641	3081	5283
In major city	0.665	-0.013	[0.037]	0.724	3081	5283
In the six months prior to the lottery:						
Employed in NZ	0.952	0.015	[0.015]	0.325	3465	6123
Log monthly NZ earnings	9.950	-0.061	[0.052]	0.241	3315	5829
Starting a new job	0.088	0.045	[0.025]	0.077	3465	6123
Pooled test that all coefficients = 0				0.349	3465	52065

Table 4 reports estimates of Equation (3), in which the dependent variable is a 'placebo' outcome which is known to be unaffected by the lottery. Standard errors are clustered at the worker level. The pooled test is conducted using a stacked regression in which lottery success and the lottery fixed effects are both interacted with indicators for each dependent variable. See Section 3 for data sources and Subsection 5.2 for details on the analytic sample and estimator.

Table 5: How winning a resident visa lottery affects job-switching and earnings

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Dependent variable indicates that a worker is employed at a firm for the first time in that month						
	Intent-to-treat; 12–36 months after lottery	Linear IV; 12–36 months after lottery	Linear IV with baseline control; 12–36 months after lottery	Linear IV; 12–48 months after lottery	Linear IV; 12–36 months after lottery; no skill-list occupations	Probit IV average partial effect; 12–36 months after lottery
Second-stage (or ITT) slope coefficient	0.010	0.015	0.014	0.013	0.017	0.014
	[0.003]	[0.004]	[0.004]	[0.004]	[0.005]	[0.003]
First-stage slope coefficient	—	0.676	0.721	0.705	0.734	—
		[0.027]	[0.024]	[0.023]	[0.026]	
First-stage F statistic	—	635.423	444.939	976.873	772.716	—
Control mean	0.019	0.019	0.019	0.019	0.019	0.019
Number of observations	149433	149433	149433	205446	93768	148806
Number of workers	3465	3465	3465	3465	2148	3459

Panel B: Dependent variable is monthly log earnings (or, in Column (6), monthly earnings)

	Intent-to-treat; 12–36 months after lottery	Linear IV; 12–36 months after lottery	Linear IV with baseline control; 12–36 months after lottery	Linear IV; 12–48 months after lottery	Linear IV; 12–36 months after lottery; no skill-list occupations	Dep. variable is earnings; Linear IV; 12–36 months after lottery
Second-stage (or ITT) slope coefficient	-0.041	-0.057	-0.025	-0.080	-0.019	-110.476
	[0.025]	[0.034]	[0.032]	[0.038]	[0.040]	[225.791]
First-stage slope coefficient	—	0.723	0.677	0.665	0.680	0.676
		[0.024]	[0.027]	[0.026]	[0.031]	[0.027]
First-stage F statistic	—	918.178	319.274	648.478	476.048	635.423
Control mean	8.406	8.406	8.408	8.424	8.350	3778.4
Number of observations	126987	126987	123771	171894	77898	149439
Number of workers	3237	3237	3153	3243	1983	3465

Table 5 presents instrumental variable (or intent-to-treat) estimates of the effect of receiving a resident visa on job-switching (Panel A) or on earnings (Panel B). Column (1) presents estimates from ordinary least squares regressions. Columns (2)–(5) (in Panel A) or (2)–(6) (in Panel B) present two-stage least squares estimates of the linear IV system (5). Column (6) in Panel A presents the average partial effect from a probit IV system which replaces the first equation of system (5) with the assumption that $y_{i,b,t} = 1\left\{\beta \text{resident}_{i,b,t} + \alpha_b + e_{i,b,t} > 0\right\}$, and additionally assumes that $e_{i,b,t}, u_{i,b,t}$ are joint normal. Column (4) is estimated using months 12–48 following the lottery, other columns are estimated using months 12–36 following the lottery. Columns (1)–(5) of Panel B additionally restrict to observations with positive earnings. Column (5) excludes entries in which the entrant held an occupation on an Essential Skills in Demand list at the time of the lottery. Standard errors are clustered at the worker level. See Section 3 for data sources and Subsection 5.2 for details on the analytic sample and estimator.

Table 6: How winning a resident visa lottery affects the earnings of those who don't switch firms

	(1)	(2)	(3)	(4)	(5)
	Lee bounds		Using the biased OLS as an upper bound	Modelling the moving decision given:	
	Without covariates	Baseline earnings as a covariate		Current earnings	Current and lagged earnings
Stayers average treatment effect	(-0.145, 0.048) [-0.266, 0.175]	(-0.135, 0.016) [-0.392, 0.141]	($-\infty$, -0.044) [$-\infty$, 0.024]	-0.050 [-0.107, 0.007]	-0.027 [-0.083, 0.029]
Number of observations	126987	123771	69327	62361	61704
Number of workers	3237	3150	2211	2982	2982

Table 6 presents the estimates of the average effect of winning a resident visa lottery on the log earnings of those workers who remain at their original firm. The specifications in columns (1), (2) and (3) yield only set-identification of the average effect; our estimates of the identified set are in parentheses while 95% confidence intervals for the identified set (which allow for clustering at the worker-level) are in brackets. The specifications in columns (4) and (5) yield point-identification; their 95% credible intervals are in brackets. Columns (1) and (2) present Lee bounds, Column (3) presents the OLS coefficient estimated only using job-stayers to form an upper-bound, while columns (4) and (5) present estimates based on a parametric model of the decision to move jobs. See Section 3 for data sources and see Subsection 5.3 for details on the analytic samples and estimators.

Table 7: The effect of including an occupation in an Essential Skills in Demand list on earnings

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Earnings measured among	All workers	Essential Skills migrants	Residents	All workers	All workers	All workers	All workers	All workers
Estimator	Two-way fixed effects	Two-way fixed effects	Two-way fixed effects	Two-way fixed effects with controls	Two-way fixed effects with controls & occ.-specific slopes	Triple-difference with Australian earnings	Proxy-confound with Australian earnings	Nonparametric average treatment effect
Effect of the Essential Skills in Demand list	0.028 [0.010]	0.038 [0.012]	0.036 [0.011]	0.037 [0.009]	0.020 [0.008]	0.019 [0.010]	0.034 [0.011]	0.025 [0.008]
Number of occ.-month obs.	161301	127320	155781	127302	127302	3567	3210	141564
Number of occupations	1029	993	1026	1020	1020	357	357	1029

Table 7 reports estimates of how including an occupation in an Essential Skills in Demand list affects log earnings in that occupation. Column (2) measures log earnings using only Essential Skills migrants, Column (3) measures log earnings using only residents and citizens, other columns measure log earnings using all workers. Columns (1), (2), (3) and (8) estimate the two-way fixed effect equation (6) using ordinary least squares. Columns (4) and (5) add controls for mean log firm value-added and for the mean log number of workers at a worker's firm, and Column (5) additionally controls for occupation-specific linear time trends. Columns (6) and (7) compare New Zealand earnings to Australian earnings; Australian data is only available at the year-by-4 digit occupation level, and so these specifications are estimated at that level (whereas other specifications are estimated at the month-by-6 digit occupation level). Column (6) presents a two-way fixed effect estimate with log New Zealand earnings minus log Australian earnings as the dependent variable. Column (7) presents the Freyaldenhoven et al. (2019) estimator, in which Australian log earnings are used as a proxy for an unobserved confound. Column (8) presents the nonparametric average treatment effect estimator discussed in Appendix D. Standard errors, clustered at the occupation level, are in parentheses. See Section 3 for data sources and Section 6 for details on the analytic sample and estimators.

Table 8: The effect of including an occupation in an Essential Skills in Demand list on intermediate outcomes

	(1)	(2)	(3)	(4)	(5)
Outcome variable	Proportion of workers employed at a firm for the first time	Log number of workers in the occupation	Mean log number of workers across workers' firms	Mean log TFP across workers' firms	Mean log value-added-per-worker across workers' firms
Effect of skill shortage listing	0.0035	0.2362	0.1033	0.0090	-0.0537
	[0.0011]	[0.0554]	[0.0540]	[0.0176]	[0.0201]
Outcome mean	0.0325	4.7268	4.3578	0.9961	11.4019
Number of occ.-month obs.	161301	161301	127521	127521	127302
Number of occupations	1029	1029	1020	1020	1020

Table 7 reports estimates of how including an occupation in an Essential Skills in Demand list affects intermediate outcomes. All specifications estimate a two-way fixed effect equation like Equation (6) (but with alternative outcome variables) using ordinary least squares. The dependent variable in Column (1) measures the proportion of workers who are observed working for a firm for the first time in that month. The dependent variable in Column (2) measures the log number of workers in the occupation. The dependent variable in Column (3) measures the mean log number of workers (in any occupation) across each worker's firm. The dependent variable in Column (4) is the mean log ratio of revenue to intermediate expenditures across each worker's firm. The dependent variable in Column (5) is the mean log value-added-per-worker across each worker's firm. See Fabling and Maré (2019) for details on the firm-level financial variables, Section 3 for other data sources and Section 6 for details on the analytic sample and estimators.

Table 9: Structural parameter estimates

	(1)	(2)	
Parameter restrictions	$\tau_{NZ} = \tau_{foreign}$; $\lambda_{NZ} = \lambda_{foreign}$	—	
Labor supply estimates		NZ-born	foreign-born
τ (labor supply elasticity for an atomistic firm)	0.814	0.845	0.751
	[0.197]	[0.208]	[0.183]
λ (workers' willingness substitute across space)	0.131	0.118	0.153
	[0.008]	[0.008]	[0.010]
p. value testing $\tau_{NZ} = \tau_{foreign}$	—	0.039	
p. value testing $\lambda_{NZ} = \lambda_{foreign}$	—	<0.001	
Production function estimates			
ρ (firms' ability to substitute across occupations)	0.919	0.919	
	[0.019]	[0.019]	
ν (returns to scale)	0.940	0.940	
	[0.014]	[0.014]	
Num. occ.-firm-demographic-year-sample observations	1205436	1205436	
Number of firms	44733	44733	
Number of occupations	951	951	

Table 9 presents parameter estimates for a structural model of wage-setting. In Column (1) we require that New Zealand-born and foreign-born workers have shared values of the structural parameters τ and λ, while in Column (2) we allow these to differ. Standard errors (which are in brackets) and hypothesis tests allow for two-way clustering within occupations and firms. See Section 3 for data sources and Section 7 for details on the model and the estimator.

A Three Models of Wage-Setting, with heterogeneous workers

In this section we extend the theoretical framework of Section 2 to allow workers to have varying roles in production. Specifically, we allow workers to have both heterogeneous (and firm-specific) productivity, and to each have an occupation (where different firms may make different use of different occupations, and occupations may be substitutes or complements). The model in this section generalizes the model in Section 2, and thus the proofs in this section serve also as proofs for the analogous results in Section 2.

We consider a static, partial-equilibrium, perfect-information labor market comprising a unit continuum of workers $\mathbf{I}$ and a finite set of firms $\mathbf{F}$. Each worker $i \in \mathbf{I}$ will be employed at a single firm $f_i \in \mathbf{F}$. Each worker $i \in \mathbf{I}$ has an (exogenous) occupation $o_i \in \mathbf{O}$ and each firm $f \in \mathbf{F}$ utilizes a set of occupations $\mathbf{O}_f \subseteq \mathbf{O}$. Let $\mathbf{F}_o \equiv \{f \in \mathbf{F} : o \in \mathbf{O}_f\}$ denote the set of firms which employ workers in occupation o.

Production. Worker $i \in \mathbf{I}$ would have productivity $l_{i,f} \in \mathbb{R}^+$, were they to work at firm f. Let $l_i \equiv \left(l_{i,f}\right)_{f \in \mathbf{F}}$. The effective labor supplied to each firm in each occupation integrates over worker productivity: $L_{f,o} \equiv \int_{i:f_i=f,o_i=o} l_{i,f} di$. Each firm f is endowed with a production function $y_f : \mathbb{R}^{|\mathbf{O}_f|} \to \mathbb{R}^+$, with which produces output using the vector of effective labor $\vec{L}_f \equiv \left(L_{f,o}\right)_{o \in \mathbf{O}_f}$. We assume that each production function y_f is differentiable and concave. We assume that output is sold at unit price in a competitive market.

Labor supply. Each worker is either a *migrant* or a *resident*. Residents can be employed at any firm while migrants can only be employed at a subset of firms. Let $\mathbf{F}_o^{\text{migrant}} \subset \mathbf{F}_o$ denote the subset of occupation-o firms at which migrants can work. Let $\mathbf{I}^{\text{migrant}} \subset \mathbf{I}$ denote the subset of workers who are migrants and let $\mathbf{I}^{\text{resident}} = \mathbf{I} \setminus \mathbf{I}^{\text{migrant}}$ denote the subset who are residents. For each worker i, let $\mathbf{F}_i \in \left\{\mathbf{F}_{o_i}^{\text{migrant}}, \mathbf{F}_{o_i}\right\}$ denote the set of firms at which she can work.

Workers have preferences over their income and over the firm which employs them. We assume that the preferences of each worker $i \in \mathbf{I}$ over their earnings W and firm f can be represented as

$$u_i(W, f) = W e^{v_{i,f}},$$

where $v_{i,f}$ measures i's taste for working at f. We assume that the distribution of $v_i \equiv \left(v_{i,f}\right)_{f \in \mathbf{F}}$ is absolutely continuous, conditional on visa status and occupation.

In each of the solution concepts specified below, a worker's earnings will be proportional to their productivity: $W_i = w_i l_{i,f_i}$. For each firm f, let $\vec{L}_f\left(\vec{w}_f; \vec{w}_{-f}\right)$ be the vector of effective labor supplied who would choose firm f, were it to pay the occupation-specific wage vector

$\vec{w}_f \equiv \left(w_{f,o}\right)_{o\in\mathbf{O}_f}$ and other firms were to pay wage vector $\vec{w}_{-f}$:

$$\vec{L}_f\left(\vec{w}_f;\vec{w}_{-f}\right) \equiv \left(L_{f,o}\left(w_{f,o};w_{-f,o}\right)\right)_{o\in\mathbf{O}_f};$$
$$L_{f,o}\left(w_{f,o};w_{-f,o}\right) \equiv L_{f,o}^{\text{resident}} + L_{f,o}^{\text{migrant}};$$
$$L_{f,o}^{\text{resident}}\left(w_{f,o};w_{-f,o}\right) \equiv \int_{i:o_i=o,f\in\arg\max_{g\in\mathbf{F}_i}\left\{u_i\left(w_{g,o_i}l_{i,f},g\right)\right\},i\in\mathbf{I}^{\text{resident}}} l_{i,f}\,di;$$
$$L_{f,o}^{\text{migrant}}\left(w_{f,o};w_{-f,o}\right) \equiv \int_{i:o_i=o,f\in\arg\max_{g\in\mathbf{F}_i}\left\{u_i\left(w_{g,o_i}l_{i,f},g\right)\right\},i\in\mathbf{I}^{\text{migrant}}} l_{i,f}\,di.$$

(By the absolute continuity of the distribution of v_i, measure zero workers will be indifferent between firms.) We assume that labor supply functions for both migrants and residents $L_{f,o}^{\text{resident}}, L_{f,o}^{\text{migrant}}$ are twice differentiable.

Defining equilibrium. We compare three solution concepts. In each solution concept we require that wages be strictly positive.

A *competitive equilibrium* comprises an assignment of workers to firms $\left(f_i\right)_{i\in\mathbf{I}}$ and a firm-occupation-specific wage schedule $\left(w_{f,o}\right)_{f\in\mathbf{F},o\in\mathbf{O}_f}$ such that each worker's firm yields her the maximal utility from her choice set:

$$\forall i\in\mathbf{I}: f_i \in \arg\max_{f\in\mathbf{F}_i} u_i\left(w_{f,o_i}l_{i,f},f\right),$$

and the mass of labor assigned to each firm maximizes its profits, taking wages as fixed:

$$\forall f\in\mathbf{F}: \vec{L}_f\left(\vec{w}_f;\vec{w}_{-f}\right) \in \arg\max_{\vec{L}\in\mathbf{R}^{|\mathbf{O}_f|}}\left\{y_f\left(\vec{L}\right) - \vec{w}_f'\cdot\vec{L}\right\}.$$

Let $\Delta_{i,f}$ denote the marginal product of worker i at firm f, given the equilibrium assignment of other workers to their firms:

$$\Delta_{i,f} = l_{i,f}\frac{\partial}{L_{f,o_i}}y_f\left(\vec{L}_f\right).$$

A *bargaining equilibrium* comprises an assignment of workers to firms $\left(f_i\right)_{i\in\mathbf{I}}$ and a worker-specific wage schedule $\left(w_i\right)_{i\in\mathbf{F}}$ such that each worker is assigned to that firm which could yield her maximal utility, were she paid her marginal product

$$\forall i\in\mathbf{I}: f_i \in \arg\max_{f\in\mathbf{F}_i} u_i\left(\Delta_{i,f},f\right),$$

and her earnings $W_i = w_i l_{i,f_i}$ are bargained *a la* Nash (1950):

$$\forall i\in\mathbf{I}: W_i \in \arg\max_{W\in\mathbf{R}^+}\left\{\left(u_i\left(W,f_i\right) - u_i^{\text{option}}\right)^{\beta}\left(\Delta_{i,f} - W\right)^{1-\beta}\right\}, \tag{26}$$

where $\beta \in (0,1)$ measures workers' bargaining power, and u_i^{option} is the maximal utility that i could receive at an alternative firm:

$$\forall i \in \mathbf{I} : u_i^{\text{option}} \equiv \max_{f \in \mathbf{F}_i \setminus \{f_i\}} u_i\left(\Delta_{i,f}, f\right).$$

A *wage-posting equilibrium* comprises an assignment of workers to firms $\left(f_i\right)_{i \in \mathbf{I}}$ and a firm-occupation-specific wage schedule $\left(w_{f,o}\right)_{f \in \mathbf{F}, o \in \mathbf{O}_f,}$ such that each worker's firm yields her the maximal utility from her choice set:

$$\forall i \in \mathbf{I} : f_i \in \underset{f \in \mathbf{F}_i}{\arg\max}\, u_i\left(w_{f,o_i} l_{i,f}, f\right),$$

and each firm's wages are chosen to maximize its profits, given the wages set by the other firms:

$$\forall f \in \mathbf{F} : \vec{w}_f \in \underset{\vec{w} \in \mathbf{R}^{|O_f|}}{\arg\max} \left\{ y_f\left(\vec{L}_f\left(\vec{w}; \vec{w}_{-f}\right)\right) - \vec{w}' \cdot \vec{L}_f\left(\vec{w}; \vec{w}_{-f}\right) \right\}.$$

We assume that there exists a unique competitive equilibrium wage schedule, a unique bargaining equilibrium wage schedule, and a unique wage-posting equilibrium wage schedule.

Relationship between this model and the model presented in Section 2. The model in Section 2 is a special case of the model in this section, corresponding to the case in which there is only a single occupation and all workers have unit productivity: $\forall i \in \mathbf{I}, f \in \mathbf{F} : l_{i,f} = 1$. In this section we explicitly assumed that labor supply functions are twice-differentiable; for the model presented in Section 2 that follows from the assumption that the random vector v_i is absolutely continuous, with a differentiable PDF.

A.1 Predictions for wage-setting

We now show that analogues to lemmas 1–3 hold in our richer environment.

Lemma A.1. *Under a competitive equilibrium, each firm $f \in \mathbf{F}$ will pay a wage vector equal to its marginal product: $\vec{w}_f = \nabla y_f\left(\vec{L}_f\right)$.*

Proof. Each firm f has the objective $\max_{\vec{L}} \left\{ y_f\left(\vec{L}\right) - \vec{w}_f' \cdot \vec{L} \right\}$. By the assumption that y_f is concave and differentiable, the objective will be concave and differentiable, and thus equilibrium labor demand will satisfy the vector-valued first order conditions

$$\nabla y_f\left(\vec{L}_f\right) - \vec{w}_f = 0.$$

□

Lemma A.2. *Under a bargaining equilibrium, each worker $i \in \mathbf{I}$ will receive wage*

$$w_i = \frac{\beta \Delta_{i,f_i} + (1-\beta)\exp\left(v_{i,f_i^{option}} - v_{i,f_i}\right)\Delta_{i,f_i^{option}}}{l_{i,f_i}}$$

where f_i^{option} is the worker's outside option:

$$f_i^{option} \in \operatorname*{argmax}_{f \in \mathbf{F}_i \setminus \{f_i\}} u_i\left(\Delta_{i,f}, f\right).$$

Proof. From Expression (26), earnings for worker i maximize

$$\left(u_i\left(W, f_i\right) - u_i^{\text{option}}\right)^{\beta}\left(\Delta_{i,f} - W\right)^{1-\beta}$$

Substituting in the utility function $u_i(W, f) = We^{v_{i,f}}$ and differentiating yields the first order condition

$$\frac{\partial}{\partial W}\left[\left(We^{v_{i,f_i}} - u_i^{\text{option}}\right)^{\beta}\left(\Delta_{i,f} - W\right)^{1-\beta}\right]_{W=W_i} \propto \frac{\beta e^{v_{i,f_i}}}{W_i e^{v_{i,f_i}} - u_i^{\text{option}}} - \frac{1-\beta}{\Delta_{i,f} - W_i} = 0.$$

Solving that equation for W_i implies that

$$W_i = \beta\Delta_{i,f} + (1-\beta)\, u_i^{\text{option}} e^{-v_{i,f_i}}.$$

By the definition of u_i^{option} and the assumed form of the utility function:

$$u_i^{\text{option}} = \max_{f \in \mathbf{F}_i \setminus \{f_i\}}\left\{\Delta_{i,f} e^{v_{i,f}}\right\} = \Delta_{i,f_i^{\text{option}}} e^{v_{i,f_i^{\text{option}}}}.$$

and by definition, $W_i = w_i l_{i,f_i}$. □

Lemma A.3. *Under a wage-posting equilibrium, the wage paid by each firm $f \in \mathbf{F}$ to occupation $o \in \mathbf{O}_f$ will be given by*

$$w_{f,o} = \left(\frac{1}{1 + \frac{1}{\eta_{f,o}}}\right)\frac{\partial y_f}{\partial L_{f,o}}$$

where $\eta_{f,o}$ is the firm-occupation-specific elasticity of labor supply:

$$\eta_{f,o} = \frac{w_{f,o}}{L_{f,o}}\frac{\partial L_{f,o}\left(w_{f,o}; w_{-f,o}\right)}{\partial w_{f,o}}.$$

Proof. Each firm f has the objective

$$\max_{\vec{w}}\left\{y_f\left(\vec{L}_f\left(\vec{w}, \vec{w}_{-f}\right)\right) - \vec{w}' \cdot \vec{L}_f\left(\vec{w}, \vec{w}_{-f}\right)\right\}.$$

By assumption, both the function $y_f(\cdot)$ and the function $\vec{L}_f(\cdot,\cdot)$ are differentiable. The objective is thus differentiable. We have additionally assumed that equilibrium wages are positive, and so the wage vector chosen by the firm will satisfy the first-order conditions

$$\frac{\partial}{\partial w_{o,f}}\left[y_f\left(\vec{L}_f\left(\vec{w},\vec{w}_{-f}\right)\right)-\vec{w}'\cdot\vec{L}_f\left(\vec{w},\vec{w}_{-f}\right)\right]\Big|_{\vec{w}=\vec{w}_f}=\frac{\partial y_f}{\partial L_{f,o}}\frac{\partial L_{f,o}}{\partial w_{f,o}}-L_{f,o}-w_{f,o}\frac{\partial L_{f,o}}{\partial w_{f,o}}=0.$$

Solving that equation for $w_{f,o}$ yields the result. □

We now show that analogues to propositions 1 and 2 also hold.

Proposition A.1. *Consider a migrant i whose bargaining equilibrium firm f_i would be unchanged, were she a resident:*

$$f_i=\underset{f\in\mathbf{F}_{o_i}^{migrant}}{\arg\max}\,u_i\left(\Delta_{i,f},f\right)\in\underset{f\in\mathbf{F}_{o_i}}{\arg\max}\,u_i\left(\Delta_{i,f},f\right).$$

Removing her from the set of migrants $\mathbf{I}^{migrant}$ would weakly increase her bargaining equilibrium wage. Moreover, her bargaining equilibrium wage would strictly increase provided that her new outside option is better than her old, i.e. provided that

$$\max_{f\in\mathbf{F}_{o_i}\setminus\{f_i\}}u_i\left(\Delta_{i,f},f\right)>\max_{f\in\mathbf{F}_{o_i}^{migrant}\setminus\{f_i\}}u_i\left(\Delta_{i,f},f\right).$$

Proof. Let $f_i^{\text{migrant option}}$ denote i's outside option as a migrant:

$$f_i^{\text{migrant option}}=\underset{f\in\mathbf{F}^{\text{migrant}}\setminus\{f_i\}}{\arg\max}\,u_i\left(\Delta_{i,f},f\right),$$

and let $f_i^{\text{resident option}}$ denote i's outside option as a resident:

$$f_i^{\text{resident option}}=\underset{f\in\mathbf{F}\setminus\{f_i\}}{\arg\max}\,u_i\left(\Delta_{i,f},f\right).$$

(Without loss of generality, assume these firms are uniquely defined.)

When $f_i^{\text{migrant option}}=f_i^{\text{resident option}}$, the result that her wage would be unchanged follows from the definition of a bargaining equilibrium (26).

Now consider the case where $f_i^{\text{migrant option}}\neq f_i^{\text{resident option}}$ and so

$$u_i\left(\Delta_{i,f_i^{\text{resident option}}},f_i^{\text{resident option}}\right)>u_i\left(\Delta_{i,f_i^{\text{migrant option}}},f_i^{\text{migrant option}}\right). \tag{27}$$

Given the assumed utility function, Inequality (27) implies that

$$\exp\left(v_{i,f_i^{\text{resident option}}}\right)\Delta_{i,f_i^{\text{resident option}}}>\exp\left(v_{i,f_i^{\text{migrant option}}}\right)\Delta_{i,f_i^{\text{migrant option}}}.$$

The result thus follows from Lemma A.2. □

As in Proposition 2, we provide an unambiguous prediction for the effect of a market-level shock under the assumption that firms' marginal products are homogeneous and constant. Specifically, we make the following assumption:

Assumption A.1. *Firms have a homogeneous production function in which each occupation has a constant marginal product:*

$$\exists y : \mathbb{R}^{|\mathbf{O}|} \to \mathbb{R}^{+} \textit{ such that } \forall f \in \mathbf{F} : y_f = y \textit{ and } \forall o \in \mathbf{O}, \vec{L}, \vec{L}' \in \mathbb{R}^{|\mathbf{O}|} : \frac{\partial}{L_o} y\left(\vec{L}\right) = \frac{\partial}{L'_o} y\left(\vec{L}'\right).$$

We modify Assumption 2 by additionally assuming that migrants and residents' productivity distributions are proportional:

Assumption A.2. *Workers' preferences and productivity is independent of their visa status, except that migrants may be more or less productive on average:*

$$\forall f \in \mathbf{F}, o \in \mathbf{O} : \exists \alpha \in \mathbb{R}^{+} \textit{ such that } \forall x \in \mathbb{R}^2 :$$
$$\mathbb{P}\left[\left(v_{i,f}, \alpha l_{i,f}\right) \leq x \,\middle|\, i \in \mathbf{I}^{migrant}, o_i = o\right] = \mathbb{P}\left[\left(v_{i,f}, l_{i,f}\right) \leq x \,\middle|\, i \in \mathbf{I}^{resident}, o_i = o\right].$$

Similarly, we modify Assumption 3 by assuming that *effective* labor supply elasticities are increasing in other firms' wages. Given the occupation-o wage schedule $\left(w_{f,o}\right)_{g\in\mathbf{F}_o}$, denote the firm-by-occupation elasticity of residents' effective labor supply as $\eta_{f,o}^{\text{resident}}\left(\left(w_{g,o}\right)_{g\in\mathbf{F}_o}\right)$:

$$\eta_{f,o}^{\text{resident}}\left(\left(w_{g,o}\right)_{g\in\mathbf{F}_o}\right) \equiv \frac{\partial}{\partial \log w_{f,o}} \log\left(L_{f,o}^{\text{resident}}\right).$$

Assumption A.3. *Each firm-by-occupation elasticity of effective resident labor supply is increasing in other firms' wages:*

$$\forall o \in \mathbf{O}, \forall \left(w_{g,o}\right)_{g\in\mathbf{F}_o}, \forall h \neq f \in \mathbf{F}_o : \frac{\partial}{\partial w_{h,o}} \eta_{f,o}^{resident}\left(\left(w_{g,o}\right)_{g\in\mathbf{F}_o}\right) > 0.$$

Our analysis of a market-level shock will use the following lemma:

Lemma A.4. *Consider a function $f : \mathbb{R} \times X \to \mathbb{R}$ where X is a partially-ordered set. Let f be decreasing in its first argument and increasing in its second, and let the equation $w = f(w, x)$ have a solution for all $x \in X$. The solution $w(x)$ of the equation $w = f(w, x)$ is increasing in x.*

Proof. Consider $x_2 > x_1$. Decompose $w(x_2) - w(x_1)$ as

$$
\begin{aligned}
w(x_2) - w(x_1) &= f(w(x_2), x_2) - f(w(x_1), x_1) \\
&= f(w(x_2), x_2) - f(w(x_1), x_2) + f(w(x_1), x_2) - f(w(x_1), x_1).
\end{aligned}
$$

Assume towards a contradiction that $w(x_2) \le w(x_1)$. Given that f is decreasing in its first element: $f(w(x_2), x_2) \ge f(w(x_1), x_2)$. But f is increasing in its second element, so $f(w(x_1), x_2) > f(w(x_1), x_1)$. Thus $f(w(x_2), x_2) - f(w(x_1), x_2) + f(w(x_1), x_2) - f(w(x_1), x_1) > 0$ and so $w(x_2) - w(x_1) > 0$. This contradicts the assumption that $w(x_2) \le w(x_1)$. □

Lemma A.4 is an implicit function theorem for monotone functions. It is very similar to existing results in monotone comparative statics (Milgrom & Shannon, 1994), though we are unaware of it having been stated explicitly. With it in hand, we can ask how expanding migrants' job options affects wages:

Proposition A.2. *Consider expanding $\mathbf{F}_o^{migrant}$ such that occupation-o migrants can work at any firm: $\mathbf{F}_o^{migrant} = \mathbf{F}_o$. Under a competitive equilibrium, Assumption A.1 implies that workers' wages will be unchanged. Under a wage-posting equilibrium, assumptions A.1, A.2 and A.3 imply that wages at every firm will increase.*

Proof. Given Assumption A.1, the marginal product of each occupation o is exogenous. The claim about competitive equilibrium follows from Lemma A.1.

We will now prove that in a Bertrand wage-posting equilibrium, when Assumptions A.1, A.2 and A.3 hold, allowing migrants to work at any firm will increase wages at every firm.

In what follows, let $\mathbf{F}_o^{\text{migrant},0}$ denote the set of firms at which migrants could work initially. Let $\eta_{f,o}^{\text{migrant}}\left(\left(w_{g,o}\right)_{g\in\mathbf{F}_o}, \mathbf{F}_o^{\text{migrant}}\right)$ denote the elasticity of migrant labor supply to the firm f, when migrants can work at firms $\mathbf{F}_o^{\text{migrant}} \in \left\{\mathbf{F}_o^{\text{migrant},0}, \mathbf{F}_o\right\}$. We continue to use $\eta_{f,o}^{\text{resident}}\left(\left(w_{g,o}\right)_{g\in\mathbf{F}_o}\right)$ to denote the elasticity of resident labor supply to firm f (conditional on wages, the labor supply of residents does not depend on $\mathbf{F}_o^{\text{migrant}}$).

Let $\eta_{f,o}\left(\left(w_{g,o}\right)_{g\in\mathbf{F}_o}, \mathbf{F}_o^{\text{migrant}}\right)$ denote the overall elasticity of labor supply to firm f. This overall elasticity is a weighted average of the group-specific elasticities:

$$
\begin{aligned}
\eta_f\left(\left(w_{g,o}\right)_{g\in\mathbf{F}_o}, \mathbf{F}_o^{\text{migrant}}\right) = {}& \sigma_{o,f}^{\text{migrant}}\left(\left(w_{g,o}\right)_{g\in\mathbf{F}_o}, \mathbf{F}_o^{\text{migrant}}\right)\eta_{f,o}^{\text{migrant}}\left(\left(w_{g,o}\right)_{g\in\mathbf{F}_o}, \mathbf{F}_o^{\text{migrant}}\right) \\
&+\sigma_{o,f}^{\text{resident}}\left(\left(w_{g,o}\right)_{g\in\mathbf{F}_o}, \mathbf{F}_o^{\text{migrant}}\right)\eta_{f,o}^{\text{resident}}\left(\left(w_{g,o}\right)_{g\in\mathbf{F}_o}\right),
\end{aligned}
\tag{28}
$$

where $\sigma_{o,f}^{\text{migrant}}\left(\left(w_{g,o}\right)_{g\in\mathbf{F}_o}, \mathbf{F}_o^{\text{migrant}}\right)$ is the (endogenous) share of effective labor supplied by migrants and $\sigma_{o,f}^{\text{resident}}\left(\left(w_{g,o}\right)_{g\in\mathbf{F}_o}, \mathbf{F}_o^{\text{migrant}}\right)$ is the share supplied by residents.

The remainder of the proof consists of 6 steps.

Step 1: For every wage schedule $\left(w_{g,o}\right)_{g\in\mathbf{F}_o}$ and firm f: $\frac{\partial}{\partial\log w_f}\eta_{f,o}^{resident}\left(\left(w_{g,o}\right)_{g\in\mathbf{F}_o}\right)<0$.
Given the assumed form for the utility function $u_i(W,f)=We^{v_{i,f}}$, if for any two firms f,g it's the case that $u_i\left(w_f l_{i,f},f\right)>u_i\left(w_g l_{i,g},g\right)$ then, for any $c>0$, it must be the case that $u_i\left(cw_f l_{i,f},f\right)>u_i\left(cw_g l_{i,g},g\right)$. It follows that

$$\forall f\in\mathbf{F}_o:\sum_{g\in\mathbf{F}_o}\frac{\partial}{\partial\log w_{g,o}}\log\left(L_{f,o}^{\text{resident}}\right)=0.$$

As such:

$$\eta_{f,o}^{\text{resident}}=-\sum_{g\in\mathbf{F}_o\setminus\{f\}}\frac{\partial}{\partial\log w_{g,o}}\log\left(L_{f,o}^{\text{resident}}\right).$$

Differentiating both sides of that equation with respect to $\log w_{f,o}$ implies that

$$\begin{aligned}\frac{\partial}{\partial\log w_{f,o}}\eta_{f,o}^{\text{resident}}&=-\frac{\partial}{\partial\log w_{f,o}}\sum_{g\in\mathbf{F}\setminus\{f\}}\frac{\partial}{\partial\log w_{g,o}}\log\left(L_{f,o}^{\text{resident}}\right)\\&=-\sum_{g\in\mathbf{F}\setminus\{f\}}\frac{\partial}{\partial\log w_{g,o}}\frac{\partial}{\partial\log w_{f,o}}\log\left(L_{f,o}^{\text{resident}}\right)\\&=-\sum_{g\in\mathbf{F}\setminus\{f\}}\frac{\partial}{\partial\log w_{g,o}}\eta_{f,o}^{\text{resident}}.\end{aligned}$$

By Assumption A.3, for each $g\neq f$: $\frac{\partial}{\partial\log w_{g,o}}\eta_{f,o}^{\text{resident}}>0$. Thus $\frac{\partial}{\partial\log w_{f,o}}\eta_{f,o}^{\text{resident}}<0$.

Step 2: When migrants can work at any firm, the elasticity of labor supply facing each firm equals the elasticity of resident labor supply: $\eta_f\left(\left(w_{g,o}\right)_{g\in\mathbf{F}_o},\mathbf{F}_o\right)=\eta_{f,o}^{resident}\left(\left(w_{g,o}\right)_{g\in\mathbf{F}_o}\right)$.
By Assumption A.2: $\eta_{f,o}^{\text{migrant}}\left(\left(w_{g,o}\right)_{g\in\mathbf{F}_o},\mathbf{F}_o\right)=\eta_{f,o}^{\text{resident}}\left(\left(w_{g,o}\right)_{g\in\mathbf{F}_o}\right)$. With Equation (28), this implies that $\eta_f\left(\left(w_{g,o}\right)_{g\in\mathbf{F}_o},\mathbf{F}_o\right)=\eta_{f,o}^{\text{resident}}\left(\left(w_{g,o}\right)_{g\in\mathbf{F}_o}\right)$.

Step 3: When migrants can work at any firm, wages are strategic complements.
By Lemma A.3, each firm f's wage solves the equation

$$w_{f,o}=\left(\frac{1}{1+\frac{1}{\eta_{f,o}\left(w_{f,o},w_{-f,o},\mathbf{F}_o^{\text{migrant}}\right)}}\right)y'_{f,o},$$

where $y'_{f,o}$ is the marginal product of occupation-o workers at firm f; by Assumption A.1 $y'_{f,o}$ is a constant. By Step 2, $\eta_{f,o}\left(w_{f,o},w_{-f,o},\mathbf{F}_o\right)=\eta_{f,o}^{\text{resident}}\left(w_{f,o},w_{-f,o}\right)$. As such, when migrants can work at any firm (i.e. when $\mathbf{F}_o^{\text{migrant}}=\mathbf{F}_o$), then :

$$w_{f,o}=\left(\frac{1}{1+\frac{1}{\eta_{f,o}^{\text{resident}}\left(w_{f,o},w_{-f,o}\right)}}\right)y'_{f,o},\tag{29}$$

By Assumption A.3, $\eta^{\text{resident}}_{f,o}$ is increasing in $w_{-f,o}$. By Step 1, $\eta^{\text{resident}}_{f,o}$ is decreasing in $w_{f,o}$. The right hand side of Equation (29) is increasing in $\eta^{\text{resident}}_{f,o}$. Thus, by Lemma A.4, $w_{f,o}$ is increasing in $w_{-f,o}$.

Step 4: Allowing migrants to work at any firm increases the best response schedules of firms which could initially employ migrants.
In this step, fix a firm $f^* \in \mathbf{F}^{\text{migrant},0}_o$. Let $\textit{effective}\left(\left(w_{g,o}\right)_{g\in\mathbf{F}_o}, \mathbf{G}\right)$ denote the vector-valued function

$$\left[\textit{effective}\left(\left(w_{g,o}\right)_{g\in\mathbf{F}_o}, \mathbf{G}\right)\right]_f = \begin{cases} w_{f,o} & \text{if } f \in \mathbf{G}; \\ 0 & \text{if } f \notin \mathbf{G}. \end{cases}$$

By Assumption A.2, migrants' preferences over firms are the same as residents would be, were wages at firms *not* in $\mathbf{F}^{\text{migrant}}$ sufficiently low that no resident would want to work there. Thus:

$$\eta^{\text{migrant}}_{f^*,o}\left(\left(w_{g,o}\right)_{g\in\mathbf{F}_o}, \mathbf{F}^{\text{migrant},0}_o\right) = \eta^{\text{resident}}_{f^*,o}\left(\textit{effective}\left(\left(w_{g,o}\right)_{g\in\mathbf{F}_o}, \mathbf{F}^{\text{migrant},0}_o\right)\right). \tag{30}$$

By construction, $\left(w_{g,o}\right)_{g\in\mathbf{F}_o} > \textit{effective}\left(\left(w_{g,o}\right)_{g\in\mathbf{F}_o}, \mathbf{F}^{\text{migrant},0}_o\right)$ (where $>$ is the standard vector partial order), and given that $f^* \in \mathbf{F}^{\text{migrant},0}_o$: $w_{f^*} = \textit{effective}\left(\left(w_{g,o}\right)_{g\in\mathbf{F}_o}, \mathbf{F}^{\text{migrant},0}_o\right)_{f^*}$. Thus, by Assumption A.3,

$$\eta^{\text{resident}}_{f^*,o}\left(\textit{effective}\left(\left(w_{g,o}\right)_{g\in\mathbf{F}_o}, \mathbf{F}^{\text{migrant},0}_o\right)\right) < \eta^{\text{resident}}_{f^*,o}\left(\left(w_{g,o}\right)_{g\in\mathbf{F}_o}\right). \tag{31}$$

Combining expressions (30) and (31) implies that

$$\eta^{\text{migrant}}_{f^*,o}\left(\left(w_{g,o}\right)_{g\in\mathbf{F}_o}, \mathbf{F}^{\text{migrant},0}_o\right) < \eta^{\text{resident}}_{f^*,o}\left(\left(w_{g,o}\right)_{g\in\mathbf{F}_o}\right) \tag{32}$$

Equation (28) implies that $\eta_{f^*,o}\left(\left(w_{g,o}\right)_{g\in\mathbf{F}_o}, \mathbf{F}^{\text{migrant},0}_o\right)$ is a convex combination of $\eta^{\text{resident}}_{f^*,o}\left(\left(w_{g,o}\right)_{g\in\mathbf{F}_o}\right)$ and of $\eta^{\text{migrant}}_{f^*,o}\left(\left(w_{g,o}\right)_{g\in\mathbf{F}_o}, \mathbf{F}^{\text{migrant},0}_o\right)$. By Step 2: $\eta_{f^*,o}\left(\left(w_{g,o}\right)_{g\in\mathbf{F}_o}, \mathbf{F}_o\right) = \eta^{\text{resident}}_{f^*,o}\left(\left(w_{g,o}\right)_{g\in\mathbf{F}_o}\right)$. Given inequality (32), it must therefore be the case that

$$\eta_{f^*,o}\left(\left(w_{g,o}\right)_{g\in\mathbf{F}_o}, \mathbf{F}_o\right) > \eta_{f^*,o}\left(\left(w_{g,o}\right)_{g\in\mathbf{F}_o}, \mathbf{F}^{\text{migrant},0}_o\right). \tag{33}$$

By Lemma A.3, firm f^*'s best-response schedule $w_{f^*,o}\left(w_{-f^*,o}; \mathbf{F}^{\text{migrant}}_o\right)$ solves the equation

$$w_{f^*,o} = \left(\frac{1}{1 + \frac{1}{\eta_{f^*,o}\left(w_{f^*,o}, w_{-f^*,o}, \mathbf{F}^{\text{migrant}}_o\right)}}\right) y'_{f^*,o}. \tag{34}$$

Treating $w_{-f^*,o}$ as fixed, Step 1 implies that the right hand side of Equation (34) is decreasing in $w_{f^*,o}$ whereas Expression (33) implies that the right hand side of Equation (34) increases when $\mathbf{F}^{\text{migrant}}_o$ expands from $\mathbf{F}^{\text{migrant},0}_o$ to $\mathbf{F}_o$. By Lemma A.4, for any $w_{-f^*,o}$:

$$w_{f^*,o}\left(w_{-f^*,o}; \mathbf{F}^{\text{migrant},0}_o\right) < w_{f^*,o}\left(w_{-f^*,o}; \mathbf{F}_o\right).$$

Step 5: Allowing migrants to work at any firm leaves unchanged the best response schedules of firms which could not initially employ migrants.

Fix a firm $f^* \notin \mathbf{F}_o^{\text{migrant},0}$. For such a firm, $\eta_{f^*,o}\left(w_{f^*,o}, w_{-f^*,o}, \mathbf{F}_o^{\text{migrant},0}\right) = \eta_{f^*,o}^{\text{resident}}\left(w_{f^*,o}, w_{-f^*,o}\right)$. By Step 2, $\eta_{f^*,o}\left(w_{f^*,o}, w_{-f^*,o}, \mathbf{F}_o\right) = \eta_{f^*,o}^{\text{resident}}\left(w_{f^*,o}, w_{-f^*,o}\right)$. Thus:

$$\eta_{f^*}\left(w_{f^*}, w_{-f^*}, \mathbf{F}_o^{\text{migrant},0}\right) = \eta_{f^*}\left(w_{f^*}, w_{-f^*}, \mathbf{F}_o\right). \tag{35}$$

By Lemma 3, each firm f's best response schedule solves the equation

$$w_{f,o} = \left(\frac{1}{1 + \frac{1}{\eta_f\left(w_{f,o}, w_{-f,o}, \mathbf{F}_o^{\text{migrant}}\right)}}\right) y'_{f,o},$$

given $w_{-f,o}$. By Equation (35) this best response schedule will be invariant over $\mathbf{F}_o^{\text{migrant}} \in \left\{\mathbf{F}_o^{\text{migrant},0}, \mathbf{F}_o\right\}$.

Step 6: Allowing migrants to work at any firm will increase wages at every firm.

Let $\left(w_{g,o}^0\right)_{g\in\mathbf{F}_o}$ denote the equilibrium wages when migrants can only work at firms in $\mathbf{F}_o^{\text{migrant},0}$, and let $\left(w_{g,o}^*\right)_{g\in\mathbf{F}_o}$ denote the equilibrium wages when migrants can work at any firm in $\mathbf{F}_o$. Let $BR\left(\left(w_{g,o}\right)_{g\in\mathbf{F}_o}, \mathbf{F}_o^{\text{migrant}}\right)$ denote the best-response schedule which returns the wage that each firm would choose were other firms to choose wage $\left(w_{g,o}\right)_{g\in\mathbf{F}_o}$ when migrants can work at firms $\mathbf{F}_o^{\text{migrant}}$. By steps 4 and 5:

$$\left(w_{g,o}^0\right)_{g\in\mathbf{F}_o} < BR\left(\left(w_{g,o}^0\right)_{g\in\mathbf{F}_o}, \mathbf{F}_o\right).$$

Let $\left(w_{g,o}^1\right)_{g\in\mathbf{F}_o} \equiv BR\left(\left(w_{g,o}^0\right)_{g\in\mathbf{F}_o}, \mathbf{F}_o\right)$, let $\left(w_{g,o}^2\right)_{g\in\mathbf{F}_o} \equiv \mathrm{BR}\left(\left(w_{g,o}^1\right)_{g\in\mathbf{F}_o}, \mathbf{F}_o\right)$, etc. By Step 3, the sequence $\left(\left(w_{g,o}^k\right)_{g\in\mathbf{F}_o}\right)_{k=0}^{\infty}$ is monotonically increasing. By Lemma A.3 the sequence is bounded above by firms' marginal products, and thus by the monotone convergence theorem the sequence converges. Given that best responses converge (and that the best response schedule is continuous), they converge to the equilibrium. Thus:

$$\left(w_{g,o}^0\right)_{g\in\mathbf{F}_o} < \left(w_{g,o}^1\right)_{g\in\mathbf{F}_o} < \left(w_{g,o}^2\right)_{g\in\mathbf{F}_o} < \ldots < \lim_{k\to\infty}\left(w_{g,o}^k\right)_{g\in\mathbf{F}_o} = \left(w_{g,o}^*\right)_{g\in\mathbf{F}_o},$$

completing the proof of Step 6 and thus of Proposition A.2. □

B Identification of the Stayers Average Treatment Effect

This appendix formalizes the identification arguments made in Subsection 5.3. To simplify the exposition, we here consider a stylized setting in which workers are observed for only two periods: once before the lottery, and once after. Our results naturally extend to the case where workers are observed for additional periods.

Consider a population Ω of workers, who are all observed in two periods $t \in \{0,1\}$. Timing is as follows:

0.1 Each worker has an initial firm.

0.2 Each worker receives earnings from their initial firm.

0.3 The lottery is held, and workers learn whether they have become residents.

1.1 Workers optionally move to a new firm.

1.2 Each worker receives earnings from their current firm.

We use the potential outcomes framework to represent workers' counterfactual outcomes: let $Y_{i,t}(W,M)$ denote the potential log earnings of worker $i \in \Omega$ in period $t \in 0,1$; log earnings depend on whether they won their lottery W and whether they moved firms M. Whether the worker moves firms will itself depend on whether they won their lottery; we represent this dependence with an additional potential outcome $M_i(W)$. As is standard in the potential outcomes framework, we abuse notation by also letting $Y_{i,t}$ and M_i represent *realized* outcomes; we thus have that

$$Y_{i,t} = Y_{i,t}\left(W_i, M_i\right) = Y_{i,t}\left(W_i, M_i\left(W_i\right)\right),$$

where W_i indicates whether i won their lottery.

We hope to identify the *stayers average treatment effect*:

$$\text{ATE}_{\text{stayers}} \equiv \mathbb{E}\left[Y_{i,1}(1,0) - Y_{i,1}(0,0) \mid M_i(1) = 0\right].$$

In words: $\text{ATE}_{\text{stayers}}$ measures the average effect of winning a resident visa lottery on workers' log earnings, among those workers who do not switch jobs.

B.1 Lee bounds

We first present weak conditions under which $\text{ATE}_{\text{stayers}}$ can be set-identified.

Assumption B.1 (random lottery). *The lottery is independent of workers' potential outcomes:*

$$W_i \perp\!\!\!\perp \left(M_i(W)\right)_{W \in 0,1}, \left(Y_{i,t}(W,M)\right)_{W,M,t \in 0,1}.$$

In our application, winning the lottery is only random conditional on the lottery a worker has entered. With Assumption B.1, we consider the simpler case in which the lottery is unconditionally random. In our empirical analysis we relax Assumption B.1 by allowing for selection into different lotteries.

Assumption B.2 (monotonicity). *If a worker would switch jobs though she lost the lottery, she also would have switched jobs had she won the lottery:*

$$\forall i \in \Omega : M_i(1) \geq M_i(0).$$

Assumption B.2 is plausible in our context, given that obtaining a resident visa expands the set of available job options.

Under assumptions B.1 and B.2, we can bound $\text{ATE}_{\text{stayers}}$ using the empirical outcome distribution. Let q denote the (observed) change in probability of moving, as a result of winning the lottery:

$$q \equiv \mathbb{P}\left[M_i(1) > M_i(0) | M_i(0) = 0\right] = \frac{\mathbb{P}\left[M_i = 1 | W_i = 1\right] - \mathbb{P}\left[M_i = 1 | W_i = 0\right]}{\mathbb{P}\left[M_i = 0 | W_i = 0\right]},$$

where the second equality follows from assumptions B.1 and B.2. To form a lower bound for $\text{ATE}_{\text{stayers}}$, consider dropping the bottom $(100 \times q)$% of job-staying lottery losers:

$$\underline{\text{ATE}_{\text{stayers}}} \equiv \mathbb{E}\left[Y_{i,1} \middle| M_i = 0; W_i = 1\right] - \mathbb{E}\left[Y_{i,1} \middle| M_i = 0; W_i = 0; Y_{i,1} \geq y_q\right],$$

where y_q solves $P\left[Y_{i,1} < y_q \middle| M_i = 0; W_i = 0\right] = q$. Similarly, we can form an upper bound for $\text{ATE}_{\text{stayers}}$, by dropping the top $(100 \times q)$% of job-staying lottery losers:

$$\overline{\text{ATE}_{\text{stayers}}} \equiv \mathbb{E}\left[Y_{i,1} \middle| M_i = 0; W_i = 1\right] - \mathbb{E}\left[Y_{i,1} \middle| M_i = 0; W_i = 0; Y_{i,1} \leq y_{1-q}\right],$$

where y_{1-q} is defined similarly

Proposition B.1. *Given assumptions B.1 and B.2, the stayers average treatment effect will lie within these bounds:*

$$\underline{ATE_{stayers}} \leq ATE_{stayers} \leq \overline{ATE_{stayers}}.$$

Proof. See Proposition 1a in Lee (2009). □

B.2 Using the OLS estimand as an upper bound

We now consider an additional assumption, which can tighten the upper bound on the stayers average treatment effect.

Assumption B.3 (winners positively selected). *Counterfactual earnings among lottery-winning job-stayers are higher than among lottery-losing job-stayers:*

$$\mathbb{E}\left[Y_{i,1}(0,0)\middle|M_i(0)=0\right] \leq \mathbb{E}\left[Y_{i,1}(0,0)\middle|M_i(1)=0\right].$$

Assumption B.3 requires that lottery-winners be disproportionately moving out of low-paying jobs.

Consider the OLS estimand which conditions on the set of observed job-stayers:

$$\beta^{\text{OLS}} \equiv \mathbb{E}\left[Y_{i,1}\middle|W_i=1;M_i=0\right] - \mathbb{E}\left[Y_{i,1}\middle|W_i=0;M_i=0\right].$$

Proposition B.2. *Given assumptions B.1 and B.3, the stayers average treatment effect will be no greater than the OLS estimand:*

$$ATE_{stayers} \leq \beta^{OLS}.$$

Proof. First, note that

$$\begin{aligned}\mathbb{E}\left[Y_{i,1}\middle|W_i=1;M_i=0\right] &= \mathbb{E}\left[Y_{i,1}(1,0)\middle|W_i=1;M_i(1)=0\right]\\ &= \mathbb{E}\left[Y_{i,1}(1,0)\middle|M_i(1)=0\right],\end{aligned}$$

where the second equality follows from Assumption B.1.

Second, note that

$$\begin{aligned}\mathbb{E}\left[Y_{i,1}\middle|W_i=0,M_i=0\right] &= \mathbb{E}\left[Y_{i,1}(0,0)\middle|W_i=0,M_{it}(0)=0\right]\\ &= \mathbb{E}\left[Y_{i,1}(0,0)\middle|M_i(0)=0\right]\\ &\leq \mathbb{E}\left[Y_{i,1}(0,0)\middle|M_i(1)=0\right],\end{aligned}$$

where the second equality follows from Assumption B.1 and the inequality follows from Assumption B.3.

Combining these results with the definitions of $\text{ATE}_{\text{stayers}}$ and β^{OLS} yields Proposition B.2. □

B.3 Modeling the moving decision

Finally, we consider assumptions which yield point identification of $\text{ATE}_{\text{stayers}}$.

Assumption B.4 (losers stay). *Lottery losers never switch jobs:* $\forall i \in \Omega : M_i(0) = 0$.

Assumption B.4 strengthens Assumption B.2. As discussed in Subsection 5.3, we rely on Assumption B.4 to prove *nonparametric* identification. In our empirical analysis we relax Assumption B.4 at the cost of additional parametric structure.

Assumption B.5 (Markovian moves). *The decision to move firms is Markovian in counterfactual earnings:*

$$M_i(1) \perp\!\!\!\perp Y_{i,0}(0,0) \mid Y_{i,1}(0,0).$$

Assumption B.5 requires that a worker's decision to move firms is independent of historic earnings, given what the worker would earn were they not to move firms. We view Assumption B.5 as plausible in contexts in which historic earnings are uninformative about the benefit of staying at a job, given current earnings. Assumption B.5 might fail if jobs differ in their earnings trajectories, or when workers' decisions to leave a job depends on their accumulated wealth.

Assumption B.6 (arrow of time). $Y_{i,0}(0,0) = Y_{i,0}(1,0) = Y_{i,0}(0,1) = Y_{i,0}(1,1)$.

Assumption B.6 requires that winning the lottery or moving firms in period 1 has no effect on period 0 outcomes.

Let both $Y_{i,0}(0,0)$ and $Y_{i,1}(0,0)$ have support $\mathcal{Y}$. We require that probability of switching given either $Y_{i,0}(0,0)$ objects are square-integrable on $\mathcal{Y}$:

Assumption B.7 (regularity). *Both the function* $\mathbb{P}\left[M_i(1) = 1 \middle| Y_{i,0}(0,0) = y_0\right]$ *and the function* $\mathbb{P}\left[M_i(1) = 1 \middle| Y_{i,1}(0,0) = y_1\right]$ *lie within* $\mathcal{H}$*, a known subset of square-integrable functions on* $\mathcal{Y}$*:*

$$\left\{\mathbb{P}\left[M_i(1) = 1 \middle| Y_{i,0}(0,0) = y_0\right], \mathbb{P}\left[M_i(1) = 1 \middle| Y_{i,1}(0,0) = y_1\right]\right\} \subset \mathcal{H} \subset \mathcal{L}^2(\mathcal{Y}),$$

while the function $d\mathbb{P}\left[Y_{i,1}(0,0) \leq y_1 \mid Y_{i,0}(0,0) = y_0\right]$ *is square-integrable on* $\mathcal{Y}^2$.

Finally, we require that the data generating process for counterfactual earnings be sufficiently non-degenerate:

Assumption B.8 (completeness). *The null space of the kernel* $d\mathbb{P}\left[Y_{i,1}(0,0) \leq y_1 \mid Y_{i,0}(0,0) = y_0\right]$ *is 0:*

$$\forall f \in \mathcal{H} : \textit{ if} \int_{y_1 \in \mathcal{Y}} f(y_1)\, dF_{Y_{i1}(0,0)}\left(y_1 \mid Y_{i,0}(0,0) = y_0\right) = 0, \textit{ then for all } y_1 : f(y_1) = 0,$$

where $\mathcal{H}$ *is the function space assumed by Assumption B.7.*

If earnings have discrete support, Assumption B.8 requires that its transition matrix be invertible. Assumption B.8 will also be satisfied whenever $Y_{i,1}(0,0)$ is increasing in $Y_{i,0}(0,0)$ (in the sense of first order stochastic dominance), as would be the case if they were positively correlated normals. Assumption B.8 is similar to the completeness condition required by the nonparametric instrumental variables literature (D'Haultfoeuille, 2011).

Proposition B.3. *Given assumptions B.1 and B.4–B.8, the stayers average treatment effect is identified from the joint distribution of earnings, moving indicators and lottery-winning indicators* $\left(Y_{i,0}, Y_{i,1}, M_i, W_i\right)_{i\in\Omega}$.

Proof. First, note that, because the lottery is random (Assumption B.1), we can infer use lottery-winning job-stayers to infer $\mathbb{E}\left[Y_{i,1}(1,0) \mid M_i(1)=0\right]$:

$$\begin{aligned}\mathbb{E}\left[Y_{i,1}(1,0) \mid M_i(1)=0\right] &= \mathbb{E}\left[Y_{i,1}(1,0)\middle|M_i(1)=0, W_i=1\right] \\ &= \mathbb{E}\left[Y_{i,1}\middle|M_i=0, W_i=1\right],\end{aligned}$$

which is observed. It thus remains to identify $\mathbb{E}\left[Y_{i,1}(0,0) \mid M_i(1)=0\right]$.

We will next show that we can infer how a worker's decision to move depends on her initial earnings by only looking at lottery winners. Given that the lottery is random (Assumption B.1), we can condition on lottery winners:

$$\mathbb{P}\left[M_i(1)=1\middle|Y_{i,0}(0,0)=y_0\right] = \mathbb{P}\left[M_i(1)=1\middle|Y_{i,0}(0,0)=y_0, W_i=1\right].$$

Given our timing assumption (Assumption B.6), $Y_{i,0} = Y_{i,0}(0,0)$, and so:

$$\mathbb{P}\left[M_i(1)=1\middle|Y_{i,0}(0,0)=y_0, W_i=1\right] = \mathbb{P}\left[M_i(1)=1\middle|Y_{i,0}=y_0, W_i=1\right].$$

Finally, among the population with $W_i = 1$, $M_i(1) = M_i$. Thus

$$\mathbb{P}\left[M_i(1)=1\middle|Y_{i,0}=y_0, W_i=1\right] = \mathbb{P}\left[M_i=1\middle|Y_{i,0}=y_0, W_i=1\right].$$

Combining equalities, we have

$$\mathbb{P}\left[M_i(1)=1\middle|Y_{i,0}(0,0)=y_0\right] = \mathbb{P}\left[M_i=1\middle|Y_{i,0}=y_0, W_i=1\right], \tag{36}$$

the right hand side of which is observed.

Similarly, we will now show that we can identify the joint distribution of $\left(Y_{i,0}(0,0), Y_{i,1}(0,0)\right)$ from lottery losers. Given that the lottery is random (Assumption B.1):

$$\mathbb{P}\left[Y_{i,1}(0,0) \le y_1 \;\middle|\; Y_{i,0}(0,0)=y_0\right] = \mathbb{P}\left[Y_{i,1}(0,0) \le y_1 \;\middle|\; Y_{i,0}(0,0)=y_0, W_i=0\right].$$

Assumption B.4 required that lottery losers never switch firms, and so having already conditioned on $W_i = 0$ we can additionally condition on $M_i = 0$:

$$\mathbb{P}\left[Y_{i,1}(0,0) \le y_1 \,\middle|\, Y_{i,0}(0,0) = y_0, W_i = 0\right] = \mathbb{P}\left[Y_{i,1}(0,0) \le y_1 \,\middle|\, Y_{i,0}(0,0) = y_0, W_i = 0, M_i = 0\right].$$

By our timing assumption (Assumption B.6) $Y_{i,0} = Y_{i,0}(0,0)$, and so:

$$\mathbb{P}\left[Y_{i,1}(0,0) \le y_1 \,\middle|\, Y_{i,0}(0,0) = y_0, W_i = 0, M_i = 0\right] = \mathbb{P}\left[Y_{i,1}(0,0) \le y_1 \,\middle|\, Y_{i,0} = y_0, W_i = 0, M_i = 0\right].$$

Finally, among with population with $W_i = 0$ and $M_i = 0$: $Y_{i,1}(0,0) = Y_{i,1}$. As such:

$$\mathbb{P}\left[Y_{i,1}(0,0) \le y_1 \,\middle|\, Y_{i,0} = y_0, W_i = 0, M_i = 0\right] = \mathbb{P}\left[Y_{i,1} \le y_1 \,\middle|\, Y_{i,0} = y_0, W_i = 0, M_i = 0\right].$$

Combining equalities, we have that

$$\mathbb{P}\left[Y_{i,1}(0,0) \le y_1 \,\middle|\, Y_{i,0}(0,0) = y_0\right] = \mathbb{P}\left[Y_{i,1} \le y_1 \,\middle|\, Y_{i,0} = y_0, W_i = 0, M_i = 0\right], \tag{37}$$

the right hand side of which is observed.

By the law of total probability:

$$\begin{aligned}
&\mathbb{P}\left[M_i(1) = 1 \middle| Y_{i,0}(0,0) = y_0\right] \\
&\quad = \int_{y_1 \in \mathcal{Y}} \mathbb{P}\left[M_i(1) = 1 \middle| Y_{i,0}(0,0) = y_0, Y_{i,1}(0,0) = y_1\right] d\mathbb{P}\left[Y_{i,1}(0,0) \le y_1 \,\middle|\, Y_{i,0}(0,0) = y_0\right].
\end{aligned} \tag{38}$$

By the Markovian moves assumption (B.5):

$$\mathbb{P}\left[M_i(1) = 1 \middle| Y_{i,0}(0,0) = y_0, Y_{i,1}(0,0) = y_1\right] = \mathbb{P}\left[M_i(1) = 1 \middle| Y_{i,1} = y_1\right]. \tag{39}$$

Substituting equations (36), (37) and (39) into Equation (38) yields the equation

$$\begin{aligned}
&\mathbb{P}\left[M_i = 1 \middle| Y_{i,0} = y_0, W_i = 1\right] \\
&\quad = \int_{y_1 \in \mathcal{Y}} \mathbb{P}\left[M_i(1) = 1 \middle| Y_{i,1}(0,0) = y_1\right] d\mathbb{P}\left[Y_{i,1} \le y_1 \,\middle|\, Y_{i,0} = y_0, W_i = 0, M_i = 0\right].
\end{aligned} \tag{40}$$

Equation (40) is a Fredholm integral equation of the first kind. Given assumptions B.7 and B.8, $\mathcal{H}$ contains at most one solution for $\mathbb{P}\left[M_i(1) = 1 \middle| Y_{i,1}(0,0) = y_1\right]$ in terms of $\mathbb{P}\left[M_i = 1 \middle| Y_{i,0} = y_0, W_i = 1\right]$ and $d\mathbb{P}\left[Y_{i,1} \le y_1 \,\middle|\, Y_{i,0} = y_0, W_i = 0, M_i = 0\right]$ (Vangel, 1992; Groetsch, 2007). Given that both the function $\mathbb{P}\left[M_i = 1 \middle| Y_{i,0} = y_0, W_i = 1\right]$ and the function $d\mathbb{P}\left[Y_{i,1} \le y_1 \,\middle|\, Y_{i,0} = y_0, W_i = 0, M_i = 0\right]$ are observed, it follows that we can identify $\mathbb{P}\left[M_i(1) = 1 \middle| Y_{i,1}(0,0) = y_1\right]$.

We will now use $\mathbb{P}\left[M_i(1) = 1 \middle| Y_{i,1}(0,0) = y_1\right]$ to identify $\mathbb{E}\left[Y_{i,1}(0,0) \mid M_i(1) = 0\right]$. By Assumption B.1:

$$\mathbb{P}\left[M_i(1) = 1\right] = \mathbb{P}\left[M_i(1) = 1 | W_i = 1\right] = \mathbb{P}\left[M_i = 1 | W_i = 1\right]. \tag{41}$$

Invoking both Assumption B.1 and Assumption B.4:

$$\mathbb{P}\left[Y_{i,1}(0,0) \le y_1\right] = \mathbb{P}\left[Y_{i,1}(0,0) \le y_1 \middle| W_i = 0\right] = \mathbb{P}\left[Y_{i,1}(0,0) \le y_1 \middle| W_i = 0, M_i = 0\right]$$
$$= \mathbb{P}\left[Y_{i,1} \le y_1 \middle| W_i = 0, M_i = 0\right]. \tag{42}$$

By Bayes' rule:

$$d\mathbb{P}\left[Y_{i,1}(0,0) \le y_1 \middle| M_i(1) = 1\right] = \mathbb{P}\left[M_i(1) = 1 \middle| Y_{i,1}(0,0) = y_1\right] \frac{d\mathbb{P}\left[Y_{i,1}(0,0) \le y_1\right]}{\mathbb{P}\left[M_i(1) = 1\right]}. \tag{43}$$

Substituting equations (41) and (42) into Equation (43) yields the equation

$$d\mathbb{P}\left[Y_{i,1}(0,0) \le y_1 \middle| M_i(1) = 1\right] = \mathbb{P}\left[M_i(1) = 1 \middle| Y_{i,1}(0,0) = y_1\right] \frac{d\mathbb{P}\left[Y_{i,1} \le y_1 \middle| W_i = 0, M_i = 0\right]}{\mathbb{P}\left[M_i = 1 | W_i = 1\right]}. \tag{44}$$

Each term on the right hand side of Equation (44) is identified, and thus its left hand side is as well. Integrating over $d\mathbb{P}\left[Y_{i,1}(0,0) \le y_1 \middle| M_i(1) = 1\right]$ yields $\mathbb{E}\left[Y_{i,1}(0,0) \mid M_i(1) = 0\right]$. We have thus identified both $\mathbb{E}\left[Y_{i,1}(1,0) \mid M_i(1) = 0\right]$ and $\mathbb{E}\left[Y_{i,1}(0,0) \mid M_i(1) = 0\right]$. Differencing these two expectations yields the stayers average treatment effect $\text{ATE}_{\text{stayers}}$.

□

C Estimation algorithm for the stayers average treatment effect

In this appendix we specify the algorithm with which we estimate the parametric model discussed in Subsection 5.3. The algorithm imposes Bayesian priors on the model parameters, and then calculates the posterior using a Gibbs sampler.

Model. For the reader's convenience, we here repeat the model from Subsection 5.3.

Associate each worker i with a lottery $l(i)$. Consider a worker i and time t, where time is measured relative to the date of the worker's lottery (i.e. $t = 0$ is the quarter in which the lottery was held). Let $Y_{i,t}(W, M)$ indicate the worker's potential log earnings, where W indicates that the worker won their lottery and M indicates that the worker has moved firms between quarter 0 and quarter t. The worker's decision to move firms also depends on whether she won the lottery, which we represent with another potential outcome $M_{i,t}(W)$. Letting W_i indicate whether the worker won their lottery, a worker's realized log earnings are thus $Y_{i,t} = Y_{i,t}\left(W_i, M_{i,t}\right) = Y_{i,t}\left(W_i, M_{i,t}(W_i)\right)$. Our target estimand is the average effect of winning the lottery, among workers who are do not move firms as a result, 1–2 years after the lottery is held:

$$\text{ATE}_{\text{stayers}} \equiv \mathbb{E}\left[Y_{i,t}(1,0) - Y_{i,t}(0,0) \middle| M_{i,t}(1) = M_{i,t}(0) = 0; t \in 4, \ldots, 11\right].$$

We impose the following assumptions:

1. $M_{i,t}(1) \geq M_{i,t}(0)$.

2. $Y_{i,-1}(1,0) = Y_{i,-1}(0,0)$.

3. $\mathbb{P}\left[M_{i,t} = 1 \mid W_i, \left(Y_{i,s}(0,0)\right)_{s \leq t}, M_{i,t-1} = 0\right] = \Phi\left(\beta_0 + \beta_1 W_i + \beta_2 Y_{i,t}(0,0) + \beta_3 W_i Y_{i,t}(0,0) + \delta_{l(i),t}\right)$;

4. $Y_{i,t}(0,0) = \alpha_0 + \alpha_1 Y_{i,t-1}(0,0) + \gamma_{l(i),t} + \epsilon_{i,t}; \quad \epsilon_{i,t} \sim N(0, \sigma_\epsilon)$;

5. $\left(\delta_{l,t}, \gamma_{l,t}\right) \sim N\left(0, \Sigma_{\delta\gamma}\right)$.

We condition on the population who have a stable firm in the quarter prior to the lottery: $M_{i,-1} = 0$. The model treats baseline counterfactual earnings $Y_{i,-1}(0,0)$ as fixed. Given that $Y_{i,-1}(1,0) = Y_{i,-1}(0,0)$ and that $M_{i,-1} = 0$ we can infer $Y_{i,-1}(0,0)$ from realized baseline earnings $Y_{i,-1}$. We estimate the model for 12 periods after the lottery (i.e. $t = 0, 1, ..., 11$).

Priors. We impose the following uninformative priors:

$$
\begin{aligned}
\beta &\sim N(0, 1000 I_4); \\
\alpha &\sim N(0, 1000 I_2); \\
\sigma_\epsilon^2 &\sim \text{inverse-gamma}(0.01, 0.01); \\
\Sigma_{\delta\gamma} &\sim \text{inverse-Wishard}(2, 0.01 I_2).
\end{aligned}
$$

Gibbs sampler. We construct a joint posterior for $\left(M_{i,t}, Y_{i,t}(0,0)\right)$ using a Gibbs sampler.[41] Following McCulloch and Rossi (1994) we model the probit 'selection equation' using a latent variable:

$$
\begin{aligned}
M_{i,t} &= \max\left\{M_{i,t-1}, 1\left\{m_{i,t} \geq 0\right\}\right\}; \\
m_{i,t} &\sim N\left(\beta_0 + \beta_1 W_i + \beta_2 Y_{i,t}(0,0) + \beta_3 W_i Y_{i,t}(0,0) + \delta_{l(i),t},\ 1\right).
\end{aligned}
$$

The state of the Gibbs sampler is the vector

$$
\theta^r \equiv \left(\left\{m_{i,t}^r\right\}, \left\{Y_{i,t}^r(0,0)\right\}, \beta^r, \alpha^r, \left\{\delta_{l,t}^r\right\}, \left\{\gamma_{l,t}^r\right\}, \Sigma_{\delta\gamma}^r, \sigma_\epsilon^r\right).
$$

In each iteration of the sampler r, each component of θ^r is drawn conditional on the other components of θ^r. Conditional posteriors for each component of θ^r are as follows:

- The posterior distribution of $m_{i,t}$ is truncated normal:

$$
m_{i,t}^r \sim \begin{cases}
N\left(\beta_0^r + \beta_1^r W_i + \beta_2^r Y_{i,t}^r(0,0) + \beta_3^r W_i Y_{i,t}^r(0,0) + \delta_{l(i),t}^r, 1\right) & \text{if } M_{i,t-1} = 1; \\
TN\left(\beta_0^r + \beta_1^r W_i + \beta_2^r Y_{i,t}^r(0,0) + \beta_3^r W_i Y_{i,t}^r(0,0) + \delta_{l(i),t}^r, 1, -\infty, 0\right) & \text{if } M_{i,t-1} = 0, M_{i,t} = 0; \\
TN\left(\beta_0^r + \beta_1^r W_i + \beta_2^r Y_{i,t}^r(0,0) + \beta_3^r W_i Y_{i,t}^r(0,0) + \delta_{l(i),t}^r, 1, 0, \infty\right) & \text{if } M_{i,t-1} = 0, M_{i,t} = 1,
\end{cases}
$$

  where $TN(a, b, c, d)$ is the truncated normal distribution with location parameter a, scale parameter b, lower bound c and upper bound d.

[41] For an overview of Gibbs sampler algorithms, including theoretical results on convergence, see Chib (2001).

- Some values of potential log earnings $Y_{i,t}(0,0)$ are known. Others are not: because the worker is a lottery winner, because the worker has moved firms, or because log earnings for the worker are unobserved (e.g. because she is not working). For each worker i, let $z_i^r \equiv \left(\left(Y_{i,t}^r(0,0)\right)_{t=0}^{11}, \left(m_{i,t}^r\right)_{t=0}^{11}\right)$ concatenate both potential earnings $Y_{i,t}^r(0,0)$ and the latent variable $m_{i,t}^r$ across all periods excluding the baseline period (which we treat as fixed). Let $known(i)$ denote the components of z_i which are known and let $unknown(i)$ denote the components of z_i which are unknown.
  Let $E\left[Y_{i,t}^r(0,0)\right]$ denote the marginal expectation of $Y_{i,t}^r(0,0)$:
  $$E\left[Y_{i,t}^r(0,0)\right] = \alpha_1^{t+1} Y_{i,-1}(0,0) + \sum_{s=0}^{t} \alpha_1^{t-s}(\gamma_{l(i),s} + \alpha_0).$$
  Let $E\left[m_{i,t}^r\right]$ denote the marginal expectation of $m_{i,t}^r$:
  $$E\left[m_{i,t}^r\right] = \beta_0 + \beta_1 W_i + \left(\beta_2 + \beta_3 W_i\right) E\left[Y_{i,t}^r(0,0)\right] + \delta_{l(i),t}.$$
  Let $E\left[z_i^r\right]$ concatenate both $E\left[\left(Y_{i,t}^r(0,0)\right)_{t=0}^{11}\right]$ and $E\left[\left(m_{i,t}^r\right)_{t=0}^{11}\right]$. Let $V\left[\left(Y_{i,t}^r(0,0)\right)_{t=0}^{11}\right]$ denote the autocovariance matrix of $Y_{i,t}^r(0,0)$:
  $$\begin{aligned} V\left[\left(Y_{i,t}^r(0,0)\right)_{t=0}^{11}\right]_{s,s} &= \sigma_\epsilon^2 \sum_{s'=0}^{s} \alpha_1^{2(s-s')}; \\ \forall s \neq s' : V\left[\left(Y_{i,t}^r(0,0)\right)_{t=0}^{11}\right]_{s,s'} &= \alpha_1^{|s-s'|} V\left[\left(Y_{i,t}^r(0,0)\right)_{t=0}^{11}\right]_{\min\{s,s'\},\min\{s,s'\}}. \end{aligned}$$
  Let $V\left[z_i^r\right]$ denote the covariance matrix of z_i^r:
  $$V\left[z_i^r\right] = \begin{bmatrix} V\left[\left(Y_{i,t}^r(0,0)\right)_{t=0}^{11}\right] & \left(\beta_2 + \beta_3 W_i\right) V\left[\left(Y_{i,t}^r(0,0)\right)_{t=0}^{11}\right] \\ \left(\beta_2 + \beta_3 W_i\right) V\left[\left(Y_{i,t}^r(0,0)\right)_{t=0}^{11}\right] & \left(\beta_2 + \beta_3 W_i\right)^2 V\left[\left(Y_{i,t}^r(0,0)\right)_{t=0}^{11}\right] + I_{11}. \end{bmatrix}$$
  The posterior distribution of unknown potential log earnings is normal, with mean
  $$E\left[z_i^r\right]_{unknown(i)} + V\left[z_i^r\right]_{unknown(i),known(i)} \left(V\left[z_i^r\right]_{known(i),known(i)}\right)^{-1} \left((z_i)_{known(i)} - E\left[z_i^r\right]_{known(i)}\right)$$
  and covariance
  $$V\left[z_i^r\right]_{unknown(i),unknown(i)} - V\left[z_i^r\right]_{unknown(i),known(i)} \left(V\left[z_i^r\right]_{known(i),known(i)}\right)^{-1} V\left[z_i^r\right]_{known(i),unknown(i)}.$$
- Let RHS_β^r be a matrix in which different rows correspond to different i,t observations (for $t \geq 0$):
  $$\left[\text{RHS}_\beta^r\right]_{(i,t),:} = \left[1, W_i, Y_{i,t}^r(0,0), W_i Y_{i,t}^r(0,0)\right].$$

Let LHS^r_β be a vector comprising elements $m^r_{i,t} - \delta^r_{l(i),t}$. The posterior distribution of β^r is normal with variance

$$\left(\text{RHS}^{r\prime}_\beta \text{RHS}^r_\beta + \frac{1}{1000} I_4\right)^{-1}$$

and mean

$$\left(\text{RHS}^{r\prime}_\beta \text{RHS}^r_\beta + \frac{1}{1000} I_4\right)^{-1} \text{RHS}^{r\prime}_\beta \text{LHS}^r_\beta.$$

- Let RHS^r_α be a matrix in which different rows correspond to different i, t observations (for $t \geq 0$):

$$\left[\text{RHS}^r_\alpha\right]_{(i,t),:} = \left[1, Y^r_{i,t-1}(0,0)\right].$$

Let LHS^r_α be a vector comprising elements $Y^r_{i,t}(0,0) - \gamma^r_{l(i),t}$. The posterior distribution of α^r is normal with variance

$$\left(\text{RHS}^{r\prime}_\alpha \text{RHS}^r_\alpha \frac{1}{\left(\sigma^r_\epsilon\right)^2} + \frac{1}{1000} I_2\right)^{-1}$$

and mean

$$\left(\text{RHS}^{r\prime}_\alpha \text{RHS}^r_\alpha \frac{1}{\left(\sigma^r_\epsilon\right)^2} + \frac{1}{1000} I_2\right)^{-1} \text{RHS}^{r\prime}_\alpha \text{LHS}^r_\alpha \frac{1}{\left(\sigma^r_\epsilon\right)^2}.$$

- Let N_l denote the number of workers in lottery l, let $\bar{\delta}^r_{l,t}$ be given by

$$\bar{\delta}^r_{l,t} = \frac{1}{N_l} \sum_{i:l(i)=l} \left[m^r_{i,t} - \beta_0 - \beta_1 W_i - \left(\beta_2 + \beta_3 W_i\right) Y^r_{i,t}(0,0)\right],$$

let $\tilde{\delta}^r_{l,t}$ be the expectation of $\delta_{l,t}$ conditional on $\gamma_{l,t}$:

$$\tilde{\delta}^r_{l,t} = \frac{\Sigma^r_{\delta\gamma_{1,2}}}{\Sigma^r_{\delta\gamma_{2,2}}} \gamma^r_{l,t}$$

and let τ^r_δ denote the precision of $\delta_{l,t}$ conditional on $\gamma_{l,t}$:

$$\tau^r_\delta = \frac{1}{\Sigma^r_{\delta\gamma_{1,1}} - \frac{\Sigma^{r\ 2}_{\delta\gamma_{1,2}}}{\Sigma^r_{\delta\gamma_{2,2}}}}.$$

The posterior distribution of each $\delta^r_{l,t}$ is normal, with mean

$$\frac{\bar{\delta}^r_{l,t} N_l + \tau^r_\delta \tilde{\delta}^r_{l,t}}{N_l + \tau^r_\delta}$$

and variance

$$\frac{1}{N_l + \tau^r_\delta}.$$

- Let $\bar{\gamma}^r_{l,t}$ be given by

$$\bar{\gamma}^r_{l,t} = \frac{1}{N_l} \sum_{i:l(i)=l} \left[Y^r_{i,t}(0,0) - \alpha_0 - \alpha_1 Y^r_{i,t-1}(0,0) \right],$$

let $\tilde{\gamma}^r_{l,t}$ be the expectation of $\gamma_{l,t}$ conditional on $\delta_{l,t}$:

$$\tilde{\gamma}^r_{l,t} = \frac{\Sigma^r_{\delta\gamma_{1,2}}}{\Sigma^r_{\delta\gamma_{1,1}}} \delta^r_{l,t}$$

and let τ^r_γ denote the precision of $\gamma_{l,t}$ conditional on $\delta_{l,t}$:

$$\tau^r_\gamma = \frac{1}{\Sigma^r_{\delta\gamma_{2,2}} - \frac{{\Sigma^r_{\delta\gamma_{1,2}}}^2}{\Sigma^r_{\delta\gamma_{1,1}}}}.$$

The posterior distribution of each $\gamma^r_{l,t}$ is normal, with mean

$$\frac{\bar{\gamma}^r_{l,t} N_l \left(\sigma^r_\epsilon\right)^{-2} + \tau^r_\gamma \tilde{\gamma}^r_{l,t}}{N_l \left(\sigma^r_\epsilon\right)^{-2} + \tau^r_\gamma}$$

and variance

$$\frac{1}{N_l \left(\sigma^r_\epsilon\right)^{-2} + \tau^r_\gamma}.$$

- Let there be L lotteries and note that we model $\delta_{l,t}$ for 12 periods (including period 0 but excluding period -1). Let $[\delta\gamma]^r$ be an $12L \times 2$ matrix with first column $\left(\delta^r_{l,t}\right)$ and second column $\left(\gamma^r_{l,t}\right)$. The posterior distribution of $\Sigma^r_{\delta\gamma}$ is inverse Wishart, with degrees of freedom $2 + 12B$ and scale matrix $(0.01 I_2 + [\delta\gamma]^{r\prime}[\delta\gamma]^r)^{-1}$.
- Let there be N worker-by-quarter observations. Let $\epsilon^r \equiv \left(Y^r_{i,t}(0,0) - \alpha^r_0 - \alpha^r_1 Y^r_{i,t-1}(0,0) - \gamma^r_{l(i),t}\right)$. The posterior distribution of $\sigma^{2^r}_\epsilon$ is inverse gamma, with shape parameter $0.01 + \frac{N}{2}$ and scale parameter $0.01 + \frac{\epsilon^{r\prime}\epsilon^r}{2}$.

We iterate the Gibbs sampler over 2000 iterations, discarding the first 200 iterations as a burn-in period. In each iteration we additionally calculate $\text{ATE}_{\text{stayers}}$ given the observed $Y_{i,t}(1,0)$ and the state $Y^r_{i,t}(0,0)$.

Convergence. Trace plots for the hyper-parameters are depicted in Appendix Figure A3. Evidently the sampler quickly converges to its equilibrium distribution, and the choice of a 200-iteration burn-in is sufficient.

D Estimating an Average Effect of Including an Occupation on an Essential Skills in Demand List

In this appendix, we present the estimator used in Section 6, which recovers an average effect of including an occupation on an Essential Skills in Demand list. To allow other research to use our estimator we will keep the exposition general; we work with a potential outcomes model similar to those used elsewhere in the difference-in-difference literature. After presenting this model, in this appendix we define our estimator and show that the estimator is unbiased for an average treatment effect.

Our approach is very similar to those proposed by De Chaisemartin and d'Haultfoeuille (2020) and De Chaisemartin and d'Haultfoeuille (2024). Our approach differs from De Chaisemartin and d'Haultfoeuille (2024) in that it exploits each event in which a unit's treatment changes, not just each unit's first such event.[42] Our approach differs from De Chaisemartin and d'Haultfoeuille (2020) in that it allows for dynamic treatment effects: we allow the outcome to depend not only whether the unit's current treatment status but also on its historic treatment status, up to a certain *a priori* lag.

The statistical model. Consider a balanced panel comprising a cross-section of units $g \in G$ and periods $t \in \mathcal{T} \subset \mathbb{N}$. Units are drawn independently from a population $\mathcal{G}$. The econometrician observes an outcome $Y_{g,t} \in \mathbb{R}$ and treatment status $Z_{g,t} \in \{0,1\}$ for each $(g,t) \in G \times \mathcal{T}$. With three assumptions —

Assumption D.1 (SUTVA). *A unit's outcome depends only on their own treatment.*

Assumption D.2 (no anticipation). *A unit's outcome depends only on their current and lagged, not on their future treatments.*

Assumption D.3 (finite relevant history). *A unit's outcome depends only on their current treatment and up to L lags of their treatment.*

— we can define unit g's potential outcome as $Y_{g,t}(\mathbf{Z}_{g,t})$, where $\mathbf{Z}_{g,t}$ is the vector $\mathbf{Z}_{g,t} \equiv (Z_{g,s})_{t-L \leq s \leq t}$. We impose the parallel trends assumption as:

Assumption D.4 (parallel trends). *Conditional on a treatment history $\mathbf{Z}_{g,t_0}$, changes in potential outcomes are mean independent of changes to treatments:*

$$\forall\, t_0 < s, \mathbf{Z} \in \{0,1\}^{L+1} : E\left[Y_{g,s}(\mathbf{Z}) - Y_{g,t_0}(\mathbf{Z}) \middle| \mathbf{Z}_{g,t_0}, \mathbf{Z}_{g,s}\right] = E\left[Y_{g,s}(\mathbf{Z}) - Y_{g,t_0}(\mathbf{Z}) \middle| \mathbf{Z}_{g,t_0}\right].$$

[42] Repeated switching is relatively common in our context. Among the 280 occupations ever included on an Essential Skills in Demand list during the period we study: 81 were always included, 142 experienced one switch, and the remaining 57 experienced at least two switches. For one example, 'Forest Scientist' was added to the Long Term Skill Shortage list in September 2010 and then removed in February 2018.

The estimator. Let $\text{treated}_t \equiv \{g \in G : Z_{g,t} = 1\}$. denote the set of units treated during period t. Let $\text{untreated}_t \equiv \{g \in G : Z_{g,t} = 0\}$ denote the set of units untreated during period t. Let $\text{unchanged}_{t_1,t_2}$ be the set of units with unchanged treatment status between period t_1 and month t_2:

$$\text{unchanged}_{t_1,t_2} = \left(\bigcap_{s=t_1}^{t_2} \text{treated}_s\right) \bigcup \left(\bigcap_{s=t_1}^{t_2} \text{untreated}_s\right).$$

We study both treating events, in which a unit is treated after being untreated for at least L periods:

$$\text{treating\_events}_t \equiv \text{treated}_t \bigcap \text{untreated}_{t-1} \bigcap \text{unchanged}_{t-L,t-1},$$

and untreating events, in which an event is untreated after being treated for at least L periods:

$$\text{untreating\_events}_t \equiv \text{untreated}_t \bigcap \text{treated}_{t-1} \bigcap \text{unchanged}_{t-L,t-1}.$$

For treating events in period t, the controls available to study outcomes in some later period $s \geq t$ are those units which were untreated in the L periods prior to t, and remained untreated in s:

$$\text{treating\_controls}_{t,s} \equiv \text{untreated}_s \bigcap \text{unchanged}_{t-L,s}.$$

For untreating events in period t, the controls available to study outcomes in some later period $s \geq t$ are those units which were treated in the L periods prior to t, and remained treated in s:

$$\text{untreating\_controls}_{t,s} \equiv \text{treated}_s \bigcap \text{unchanged}_{t-L,s}.$$

For each treating event, we can calculate the difference-in-difference estimate of that event using an appropriate control

$$\forall s \geq t,\ g \in \text{treating\_events}_t,\ c \in \text{treating\_controls}_{t,s} : \beta_{g,t,s,c} \equiv \left(Y_{g,s} - Y_{g,t-1}\right) - \left(Y_{c,s} - Y_{c,t-1}\right).$$

Estimates for each untreating event are similar, although we multiply them by -1 to preserve the convention that they represent the effect of being treated:

$$\forall s \geq t, g \in \text{untreating\_events}_t, c \in \text{untreating\_controls}_{t,s} : \beta_{g,t,s,c} \equiv \left(Y_{c,s} - Y_{c,t-1}\right) - \left(Y_{g,s} - Y_{g,t-1}\right).$$

In aggregating these treatment-by-month-by-control estimates, we average first across all available controls, for each event and outcome periods

$$\forall s \geq t,\ g \in \text{treating\_events}_t : \beta_{g,t,s} \equiv \frac{1}{\left|\text{treating\_controls}_{t,s}\right|} \sum_{c \in \text{treating\_controls}_{t,s}} \beta_{g,t,s,c}\,;$$

$$\forall s \geq t,\ g \in \text{untreating\_events}_t : \beta_{g,t,s} \equiv \frac{1}{\left|\text{untreating\_controls}_{t,s}\right|} \sum_{c \in \text{untreating\_controls}_{t,s}} \beta_{g,t,s,c}\,.$$

We next form an estimate for each event, averaging up to T periods after the event (we use fewer months if the unit switched back to its prior status within T periods):

$$\forall g \in \text{treating\_events}_t : \beta_{g,t} \equiv \frac{1}{S_{g,t}} \sum_{s=0}^{S_{g,t}} \beta_{g,t,s};\ S_{g,t} \equiv \min\left\{t+T, \min\{s>t : g \in \text{untreated}_s\}\right\};$$
$$\forall g \in \text{untreating\_events}_t : \beta_{g,t} \equiv \frac{1}{S_{g,t}} \sum_{s=0}^{S_{g,t}} \beta_{g,t,s};\ S_{g,t} \equiv \min\left\{t+T, \min\{s>t : g \in \text{treated}_s\}\right\}.$$

Finally, our point estimate for the average treatment effect averages across event-specific estimates:

$$\beta^{\text{ATE}} \equiv \frac{\left(\sum_t \sum_{g \in \text{treating\_events}_t} \beta_{g,t}\right) + \left(\sum_t \sum_{g \in \text{untreating\_events}_t} \beta_{g,t}\right)}{\left|\left(\bigcup_t \text{treating\_events}_t\right) \cup \left(\bigcup_t \text{untreating\_events}_t\right)\right|}.$$

Unbiasedness. An estimator β is an unbiased estimator of an average treatment effect if there exists hypothetical treatments $\left(\mathbf{z}_{g,t}\right)_{g \in \mathcal{G}, t \in \mathcal{T}}$ and $\left(\mathbf{z}'_{g,t}\right)_{g \in \mathcal{G}, t \in \mathcal{T}}$ such that $\mathbf{z}_{g,t} > \mathbf{z}'_{g,t}$ and that

$$E[\beta] = E_\mu\left[Y_{g,t}\left(\mathbf{z}_{g,t}\right) - Y_{g,t}\left(\mathbf{z}'_{g,t}\right)\right],$$

where $>$ is the strict partial order on vectors,[43] and μ is a probability measure on $\mathcal{G} \times \mathcal{T}$.

Proposition D.1. *The estimator β^{ATE} is an unbiased estimator of an average treatment effect.*

Proof. For some periods t, s with $t < s$: consider a treated unit $g \in \text{treating\_events}_t$ and a control unit $c \in \text{treating\_controls}_{t,s}$. The event-, control- and period-specific estimate $\beta_{g,t,s,c}$ can be decomposed as

$$\begin{aligned}
\beta_{g,t,s,c} &\equiv \left(Y_{g,s} - Y_{o,t-1}\right) - \left(Y_{c,s} - Y_{c,t-1}\right) \\
&= \left(Y_{g,s}\left(\mathbf{z}_{g,s}\right) - Y_{g,t-1}\left(\mathbf{z}_{g,t-1}\right)\right) - \left(Y_{c,s}\left(\mathbf{z}_{c,s}\right) - Y_{c,t-1}\left(\mathbf{z}_{c,t-1}\right)\right) \\
&= \left(Y_{g,s}\left(\mathbf{z}_{g,s}\right) - Y_{g,t-1}\left(\mathbf{z}_{g,t-1}\right)\right) - \left(Y_{c,s}\left(\mathbf{z}_{c,s}\right) - Y_{c,t-1}\left(\mathbf{z}_{c,t-1}\right)\right) - Y_{g,s}\left(\mathbf{z}_{c,s}\right) + Y_{g,s}\left(\mathbf{z}_{c,s}\right) \\
&= \left(Y_{g,s}\left(\mathbf{z}_{g,s}\right) - Y_{g,s}\left(\mathbf{z}_{c,s}\right)\right) + \left(Y_{g,s}\left(\mathbf{z}_{c,s}\right) - Y_{g,t-1}\left(\mathbf{z}_{g,t-1}\right)\right) - \left(Y_{c,s}\left(\mathbf{z}_{c,s}\right) - Y_{c,t-1}\left(\mathbf{z}_{c,t-1}\right)\right).
\end{aligned}$$

By the definition of the sets treating_events_t and $\text{treating\_controls}_{t,s}$: $\mathbf{z}_{c,s} = \mathbf{z}_{c,t-1} = \mathbf{z}_{g,t-1} = \mathbf{0} \equiv (0, 0, \ldots, 0)$. Thus:

$$\beta_{g,t,s,c} = \left(Y_{g,s}\left(\mathbf{z}_{g,s}\right) - Y_{g,s}\left(\mathbf{z}_{c,s}\right)\right) + \left(Y_{g,s}(\mathbf{0}) - Y_{g,t-1}(\mathbf{0})\right) - \left(Y_{c,s}(\mathbf{0}) - Y_{c,t-1}(\mathbf{0})\right). \tag{45}$$

Taking an expectation over both sides of Equation (45), conditioning on $g \in \text{treating\_events}_t$

[43] I.e. $\mathbf{x} > \mathbf{x}'$ iff $\mathbf{x} \neq \mathbf{x}'$ and $\forall k : \mathbf{x}_k \geq \mathbf{x}'_k$.

and $c \in \text{treating\_controls}_{t,s}$, yields

$$\begin{aligned}
&E\left[\beta_{g,t,s,c} \middle| g \in \text{treating_events}_t, c \in \text{treating_controls}_{t,s}\right] \\
=&E\Big[\left(Y_{g,s}\left(\mathbf{z}_{g,s}\right) - Y_{g,s}\left(\mathbf{z}_{c,s}\right)\right) + \left(Y_{g,s}\left(\mathbf{0}\right) - Y_{g,t-1}\left(\mathbf{0}\right)\right) - \left(Y_{c,s}\left(\mathbf{0}\right) - Y_{c,t-1}\left(\mathbf{0}\right)\right) \\
&\qquad\qquad\qquad\qquad\qquad\qquad \Big| g \in \text{treating_events}_t, c \in \text{treating_controls}_{t,s}\Big] \\
=&E\left[Y_{g,s}\left(\mathbf{z}_{g,s}\right) - Y_{g,s}\left(\mathbf{z}_{c,s}\right) \middle| g \in \text{treating_events}_t, c \in \text{treating_controls}_{t,s}\right] \\
+&E\left[Y_{g,s}\left(\mathbf{0}\right) - Y_{g,t-1}\left(\mathbf{0}\right) \middle| g \in \text{treating_events}_t, c \in \text{treating_controls}_{t,s}\right] \\
-&E\left[Y_{c,s}\left(\mathbf{0}\right) - Y_{c,t-1}\left(\mathbf{0}\right) \middle| g \in \text{treating_events}_t, c \in \text{treating_controls}_{t,s}\right]. \qquad (46)
\end{aligned}$$

By the assumption that units are IID:

$$\begin{aligned}
&E\left[Y_{g,s}\left(\mathbf{0}\right) - Y_{g,t-1}\left(\mathbf{0}\right) \middle| g \in \text{treating_events}_t, c \in \text{treating_controls}_{t,s}\right] \\
=&E\left[Y_{g,s}\left(\mathbf{0}\right) - Y_{g,t-1}\left(\mathbf{0}\right) \middle| g \in \text{treating_events}_t\right]. \qquad (47)
\end{aligned}$$

(Note the expectation on the left hand side is properly understood as an expectation over pairs, whereas the expectation on the right hand side is an expectation over units.) Note that the set treating_events_t is equal to

$$\text{treating\_events}_t = \left\{g \in G : \mathbf{Z}_{g,t-1} = \mathbf{0}, Z_{g,t} = 1\right\}.$$

It follows from Assumption D.4 that

$$E\left[Y_{g,s}\left(\mathbf{0}\right) - Y_{g,t-1}\left(\mathbf{0}\right) \middle| g \in \text{treating\_events}_t\right] = E\left[Y_{g,s}\left(\mathbf{0}\right) - Y_{g,t-1}\left(\mathbf{0}\right) \middle| \mathbf{Z}_{g,t-1} = \mathbf{0}\right]. \qquad (48)$$

Combining equations (47) and (48) yields

$$E\left[Y_{g,s}\left(\mathbf{0}\right) - Y_{g,t-1}\left(\mathbf{0}\right) \middle| g \in \text{treating\_events}_t, c \in \text{treating\_controls}_{t,s}\right] = E\left[Y_{g,s}\left(\mathbf{0}\right) - Y_{g,t-1}\left(\mathbf{0}\right) \middle| \mathbf{Z}_{g,t-1} = \mathbf{0}\right].$$

Similarly, one can show that

$$E\left[Y_{c,s}\left(\mathbf{0}\right) - Y_{c,t-1}\left(\mathbf{0}\right) \middle| g \in \text{treating\_events}_t, c \in \text{treating\_controls}_{t,s}\right] = E\left[Y_{c,s}\left(\mathbf{0}\right) - Y_{c,t-1}\left(\mathbf{0}\right) \middle| \mathbf{Z}_{c,t-1} = \mathbf{0}\right].$$

Returning to Equation (46), we see that these terms cancel and are left with

$$\begin{aligned}
&E\left[\beta_{g,t,s,c} \middle| g \in \text{treating_events}_t, c \in \text{treating_controls}_{t,s}\right] \\
&\qquad = E\left[Y_{g,s}\left(\mathbf{z}_{g,s}\right) - Y_{g,s}\left(\mathbf{z}_{c,s}\right) \middle| g \in \text{treating_events}_t, c \in \text{treating_controls}_{t,s}\right].
\end{aligned}$$

One can prove a similar result for untreating events. Proposition D.1 then follows from the linearity of unbiasedness and of sample averages. □

E Additional Details for Structural Modelling

E.1 Sample construction.

We estimate our model using a firm-occupation-year-demographic-sample-level dataset. We include financial years 2013-2021.[44] (We exclude years before the 2013 Census, when occupation would be missing for most workers.)

As elsewhere, we measure earnings using administrative tax data, which is observed at the worker-firm-month level. Our model assumes that the wage earned by a worker, given their effective labor, their occupation and their firm, is unaffected by their visa status and demographic type. To reduce measurement error in our wage estimates we calculate them using only worker-month observations with a unique associated firm. We also require that the worker reported working at least 30 hours per week in the prior census. For each worker-firm-year, we sum earnings across such months, and calculate their wage as their overall earnings divided by (number of months × weekly hours of work reported in the prior census). We then calculate the log wage $\log \hat{w}^s_{f,o,t}$ by taking the mean log wage across workers in sample s.

We treat wage-bills as measured without error; we measure the wage-bill of a firm-occupation-year observation $wL_{f,o,t}$ by aggregating over all worker-firm-month observations (i.e. regardless of the worker's random sample, and not restricting to workers with a unique firm).

Our sample-s measurement of the effective labor used by a firm in a given occupation $\hat{L}^s_{f,o,t}$ is constructed by dividing the wage-bill by the sample-s measurement of the wage: $\hat{L}^s_{f,o,t} = \frac{wL_{f,o,t}}{\hat{w}^s_{f,o,t}}$. Note that because wage bills are measured using all workers, the only source of error in $\hat{L}^s_{f,o,t}$ is due to the measurement error in wages.

Worker shares. The set of workers with which we define the market shares $\sigma^{status}_{f,o,x,t|c_f}$, $\sigma^{status}_{c,o,x,t}$ and within-job shares $\sigma_{x,status|f,o,t}$ differs from the set of workers with which we define wages: we define shares by associating each employee only with the firm that paid them the maximal earnings during each year. Note that worker shares are treated as measured without error and so are constructed to be constant across samples.

Visa status. We assume that workers who held temporary visas other than the Essential Skills visa had unrestricted job options and so we treat them as residents. This assumption is correct for the most common temporary visas other than the Essential Skills visa (the Partner of a Worker Work Visa and working holiday visas), but is incorrect for some more unusual visas (such as the Thai Chefs Work Visa, which requires that its holder work as a Thai chef).

Demographics We define a worker's type x_i as indicating whether they were born in New Zealand, which we measure using the 2013 and 2018 population censuses.

[44] New Zealand financial years end in March; the 2013 financial year for example covers April 2013-March 2014.

Firms and location. Location is defined using Statistics NZ's 'functional urban areas'. We define a 'firm' as a functional urban area × enterprise. (As such, each firm will by construction have a unique location.)

Sample restrictions. We exclude a small number of occupation-firm-year combinations with a weighted-average market-share of 1 — i.e. those for whom

$$\sum_{x,status} \left(\sigma_{x,status|f,o,t}\sigma^{status}_{c,o,x,t}\right) = 1.$$

Our model predicts that log markdowns for such observations will be $-\infty$.

We exclude observations for which the estimated wage is equal to or less than the minimum wage in that year, for which our procedure for inferring labor-augmenting productivity (which requires that we can invert the first order condition for wages) would be invalid. Extending our method to account for such observations (using moment inequality methods) would be valuable, especially for applications that directly involve studying the effects of wage floors.

Finally, when calculating our estimating equations we retain only observations for which wages are nonmissing (e.g. because workers in that sample are observed) for that observation, for both the current period and for two lags, and for the corresponding observation in the other split-sample for two lags. This ensures that our estimating equations are all calculated using a consistent set of observations. Note we need not (and do not) make this restriction when calculating the firm-level labor aggregate $\tilde{L}_{f,t}$.

E.2 Analytic derivations

The wage equation. We here derive the wage equation (13). Combining equations (11) and (12), taking logs, and solving for labor-augmenting productivity $\phi_{f,o,t}$ yields

$$\phi_{f,o,t} = \log w_{f,o,t} - \log \nu + C_{f,t} + (1-\rho)\log L_{f,o,t} + \log\left(1 + \frac{1}{\eta_{f,o,t}}\right), \tag{49}$$

where the firm-year term $C_{f,t}$ is given by

$$C_{f,t} \equiv \left(\frac{\rho-\nu}{\rho}\right)\log\left(\sum_{o'\in\mathbf{O}_{f,t}} e^{\phi_{f,o',t}} L^{\rho}_{f,o',t}\right). \tag{50}$$

Substituting Equation (49) into Equation (50), and solving for $C_{f,t}$ yields

$$C_{f,t} = \left(\frac{\rho-\nu}{\nu}\right)\log\left(\sum_{o'\in\mathbf{O}_{f,t}} w_{f,o,t} L_{f,o,t}\left(1+\frac{1}{\eta_{f,o,t}}\right)\right) + \left(\frac{\nu-\rho}{\nu}\right)\log \nu. \tag{51}$$

Substituting Equation (51) into Equation (49) and then solving for $\log w_{f,o,t}$ yields Equation (13).

The amenity value of employment. To infer the within-market amenity terms $\xi_{f,o,x,t}$ we combine equations (15) and (17) to derive

$$\xi_{f,o,x,t} = \log \sigma^{resident}_{f,o,x,t|c_f} - \tau \log(w_{f,o,t}) + D_{c_f,o,x,t}, \tag{52}$$

where $D_{c,o,x,t}$ is a location-occupation-demographic-year constant. Imposing the constraint that $\log \sum_{f \in \mathbf{F}_{c,o,t}} \exp \xi_{f,o,x,t} = 0$ on Equation (52) implies that

$$\xi_{f,o,x,t} = \log \sigma^{resident}_{f,o,x,t|c_f} - \tau \log(w_{f,o,t}) - \log \sum_{g \in \mathbf{F}_{c_f,o,t}} \exp\left(\log \sigma^{resident}_{g,o,x,t|c_f} - \tau \log(w_{g,o,t})\right). \tag{53}$$

Similarly, to infer the between-market terms $\bar{\xi}_{c,o,x,t}$ we combine equations (16) and (18) to derive

$$\bar{\xi}_{c,o,x,t} = \log \sigma^{resident}_{c,o,x,t} - \lambda \log\left(\sum_{f \in \mathbf{F}_{c,o,t}} \exp\left(\tau \log(w_{f,o,t}) + \xi_{f,o,x,t}\right)\right) + D_{o,x,t}, \tag{54}$$

where $D_{o,x,t}$ is an occupation-demographic-year constant. Imposing the constraint that $\log \sum_{c \in \mathbf{C}} \exp \bar{\xi}_{c,o,x,t} = 0$ on Equation (54) implies that

$$\begin{aligned} \bar{\xi}_{c,o,x,t} = {} & \log \sigma^{\text{resident}}_{c,o,x,t} - \lambda \log\left(\sum_{f \in \mathbf{F}_{c,o,t}} \exp\left(\tau \log(w_{f,o,t}) + \xi_{f,o,x,t}\right)\right) \\ & - \log \sum_{c' \in \mathbf{C}} \exp\left(\log \sigma^{\text{resident}}_{c',o,x,t} - \lambda \log\left(\sum_{f \in \mathbf{F}_{c',o,t}} \exp\left(\tau \log(w_{f,o,t}) + \xi_{f,o,x,t}\right)\right)\right). \end{aligned} \tag{55}$$

E.3 A GMM representation of our estimator

In this subsection we present a GMM representation of our estimator. This representation yields a numerically identical estimate to that presented in Subsection 7.3. It would be somewhat more difficult to calculate because it requires optimizing over all 6 parameters, rather than just the 4 labor supply parameters. However it facilitates standard inference. Specifically, we present standard errors that allow clustering across both firms and occupations, calculated using the approach given by Cameron, Gelbach, and Miller (2011).

Given a candidate parameter vector $(\tau_x)_{x \in \mathbf{X}}, (\lambda_x)_{x \in \mathbf{X}}, \rho, \nu$, we calculate the empirical moments as follows:

1. Within each sample s we calculate the location and firm amenity values $\hat{\bar{\xi}}^s_{c,o,x,t}, \hat{\xi}^s_{f,o,x,t}$ using equations (53) and (55), and using the wage estimate for that sample.

2. We calculate demographic-by-visa status labor supply elasticities $\eta^{status}_{f,o,x,t}$ using Equation (19)

3. We calculate the overall labor supply elasticities $\eta_{f,o,t}$ using Equation (20).

4. Given these elasticities we calculate labor-augmenting productivity $\hat{\phi}^s_{f,o,t}$ in each sample s by inverting Equation (13), using the wage estimate and the labor utilization estimate for that sample.

5. When estimating the model in which the two demographic types have the same structural labor supply parameters τ, λ, we define our moment vectors as

$$
\begin{aligned}
m^1_{f,o,x,t,s} &\equiv \left(\Delta\hat{\xi}^s_{f,o,x,t} - E\left[\Delta\hat{\xi}^s_{f,o,x,t}\middle|c_f, o, x, t, s\right]\right)\left(\Delta\hat{\phi}^{-s}_{f,o,t} - E\left[\Delta\hat{\phi}^{-s}_{f,o,t}\middle|c_f, o, x, t, -s\right]\right),\\
m^2_{f,o,t,x,s} &\equiv \left(\Delta\sum_{o\in\mathbf{O}_{f,t}} e^{\mathcal{L}^{*s}_{f,o,x,t}} - E\left[\Delta\sum_{o\in\mathbf{O}_{f,t}} e^{\mathcal{L}^{*s}_{f,o,x,t}}\middle|c_f, o, x, t, s\right]\right)\left(\Delta\hat{\phi}^{-s}_{f,o,t} - E\left[\Delta\hat{\phi}^{-s}_{f,o,t}\middle|c_f, o, x, t, -s\right]\right),\\
m^3_{f,o,x,t,s} &\equiv \left(\Delta\hat{\xi}^s_{f,o,x,t} - E\left[\Delta\hat{\xi}^s_{f,o,x,t}\middle|c_f, o, x, t, s\right]\right)\left(\Delta\hat{\phi}^{-s}_{f,o,t-1} - E\left[\Delta\hat{\phi}^{-s}_{f,o,t-1}\middle|c_f, o, x, t, -s\right]\right),\\
m^4_{f,o,x,t,s} &\equiv \left(\Delta\hat{\hat{\xi}}^s_{c,o,x,t}, -E\left[\hat{\hat{\xi}}^s_{c,o,x,t},\middle|o, x, t, s\right]\right)\left(\Delta\log\left(\sum_{f\in\mathbf{F}_{c,o,t}}\exp\left(\tau^0_x\log w^{*-s}_{f,o,t}\right)\right)\right.\\
&\qquad\left. -E\left[\Delta\log\left(\sum_{f\in\mathbf{F}_{c,o,t}}\exp\left(\tau^0_x\log w^{*-s}_{f,o,t}\right)\right)\middle|o, x, t, -s\right]\right),
\end{aligned}
$$

where $E[\cdot|\cdot]$ is the empirical conditional mean, $\mathcal{L}^{*s}_{f,o,x,t}$ is defined as in Step 5 of the procedure given in Subsection 7.3, and $w^{*s}_{f,o,t}$ is defined as in Step 8 of the procedure given in Subsection 7.3.

When estimating the model in which the two demographic types are allowed to have differing values of the structural labor supply parameters, $m^1_{f,o,x,t,s}$ and $m^2_{f,o,x,t,s}$ are defined as above. However, instead of $m^3_{f,o,t,x,s}$ and $m^4_{f,o,t,x,s}$, we have

$$
m^{3,NZ}_{f,o,x,t,s} \equiv \begin{cases} m^3_{f,o,x,t,s} & \text{if } x = NZ;\\ 0 & \text{otherwise,}\end{cases}
$$

$$
m^{3,foreign}_{f,o,x,t,s} \equiv \begin{cases} m^3_{f,o,x,t,s} & \text{if } x = foreign;\\ 0 & \text{otherwise,}\end{cases}
$$

$$
m^{4,NZ}_{f,o,x,t,s} \equiv \begin{cases} m^4_{f,o,x,t,s} & \text{if } x = NZ;\\ 0 & \text{otherwise,}\end{cases}
$$

$$
m^{4,foreign}_{f,o,x,t,s} \equiv \begin{cases} m^4_{f,o,x,t,s} & \text{if } x = foreign;\\ 0 & \text{otherwise.}\end{cases}
$$

E.4 Imputation of amenities when no types are observed at a firm

When estimating our structural parameters we omit firm-year-demographic observations at which no residents are observed. However, when calculating counterfactuals we include such observations. To do so, we must impute the amenity value of employment. We impute these amenities by approximating the posterior mode under the assumption that amenities are joint-normally distributed across types.

Specifically, we first calculate the empirical covariance of type-*NZ* and type-*foreign* amenities across firm-by-occupations at which both are observed. We then use this covariance and the observed amenities to form a normal prior for unobserved amenities; let $\mu^{\xi}_{f,o,x,t}$ and $s^{\xi}_{f,o,x,t}$ denote the mean and standard deviation of this prior. Equations (15) and (17) imply that, given the vector of amenities $\tilde{\xi}_{f,o,x,t}$, the distribution of occupation-o type-x residents across firms in location c will be multinomial with assignment probability $\frac{\exp\left(\tau\log(w_{f,o,t})+\tilde{\xi}_{f,o,x,t}\right)}{\sum_{f'\in\mathbf{F}_{c,o,t}}\exp\left(\tau\log(w_{f',o,t})+\tilde{\xi}_{f',o,x,t}\right)}$. The log posterior for $\tilde{\xi}_{f,o,x,t}$, given that no type-x occupation-o residents are employed at firm f, is thus equal to

$$C+n_{x,c,t}\log\left(\frac{\sum_{f'\in\mathbf{F}_{c,o,t}\setminus f}\exp\left(\tau\log(w_{f',o,t})+\tilde{\xi}_{f',o,x,t}\right)}{\sum_{f'\in\mathbf{F}_{c,o,t}}\exp\left(\tau\log(w_{f',o,t})+\tilde{\xi}_{f',o,x,t}\right)}\right)-\frac{\left(\tilde{\xi}_{f,o,x,t}-\mu^{\xi}_{f,o,x,t}\right)^2}{2{s^{\xi}_{f,o,x,t}}^2},$$

where the constant C does not depend on the amenities and $n_{x,c,t}$ is the number of type-x workers in location c during period t. Taking a first-order condition for $\tilde{\xi}_{f,o,x,t}$ implies that the posterior mode solves the vector equation

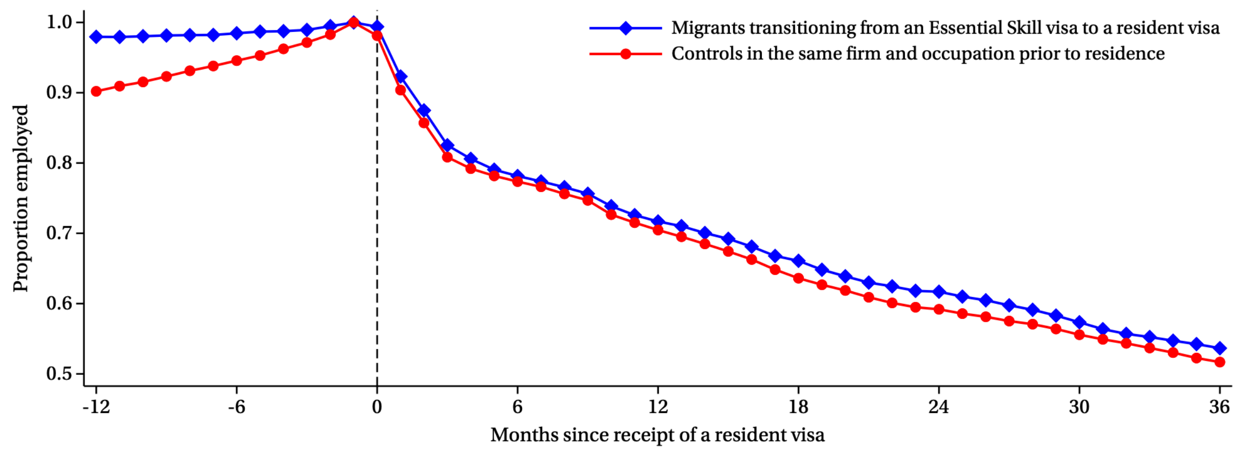
$$-n_{x,c,t}\left(\frac{\exp\left(\tau\log(w_{f,o,t})+\tilde{\xi}_{f,o,x,t}\right)}{\sum_{f'\in\mathbf{F}_{c,o,t}}\exp\left(\tau\log(w_{f',o,t})+\tilde{\xi}_{f',o,x,t}\right)}\right)-\frac{\left(\tilde{\xi}_{f,o,x,t}-\mu^{\xi}_{f,o,x,t}\right)}{{s^{\xi}_{f,o,x,t}}^2}=0. \tag{56}$$

This equation lacks a closed-form solution. We find an approximate solution by replacing the term $\left(\frac{\exp\left(\tau\log(w_{f,o,t})+\tilde{\xi}_{f,o,x,t}\right)}{\sum_{f'\in\mathbf{F}_{c,o,t}}\exp\left(\tau\log(w_{f',o,t})+\tilde{\xi}_{f',o,x,t}\right)}\right)$ with its evaluation at the prior $\tilde{\xi}_{f',o,x,t}\approx\mu^{\xi}_{f',o,x,t}$. This approximation simplifies Equation (56) to a linear equation, yielding our imputed value of $\tilde{\xi}_{f,o,x,t}$ as

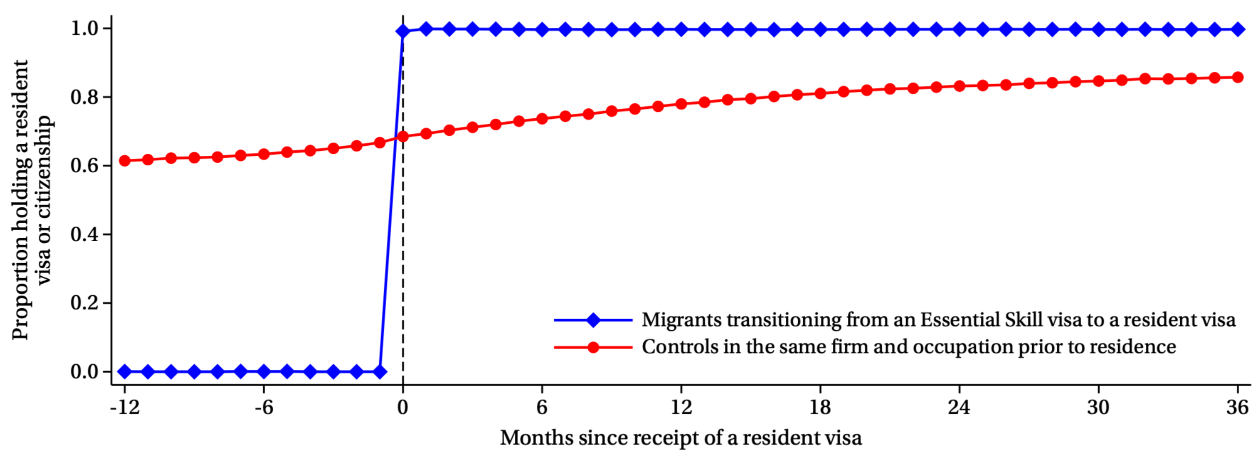
$$\tilde{\xi}_{f,o,x,t}=\mu^{\xi}_{f,o,x,t}-{s^{\xi}_{f,o,x,t}}^2 n_{x,c,t}\left(\frac{\exp\left(\tau\log(w_{f,o,t})+\mu^{\xi}_{f,o,x,t}\right)}{\sum_{f'\in\mathbf{F}_{c,o,t}}\exp\left(\tau\log(w_{f',o,t})+\mu^{\xi}_{f',o,x,t}\right)}\right).$$

We impute amenities for occupation-by-location markets at which no type-x workers are observed similarly.

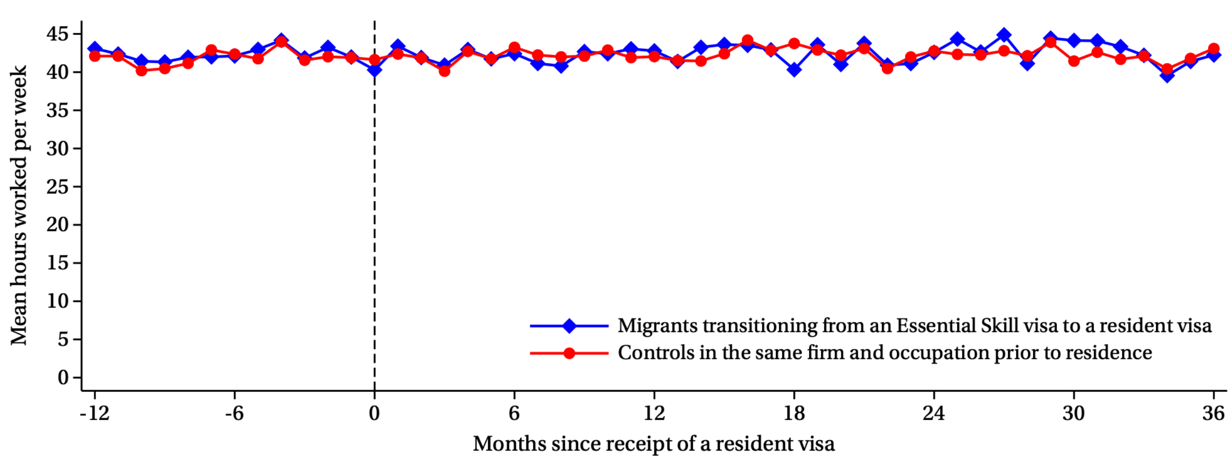
Figure A1: How employment, visa status, hours of work and firm size change when a migrant receives a resident visa



Panel A: Employment

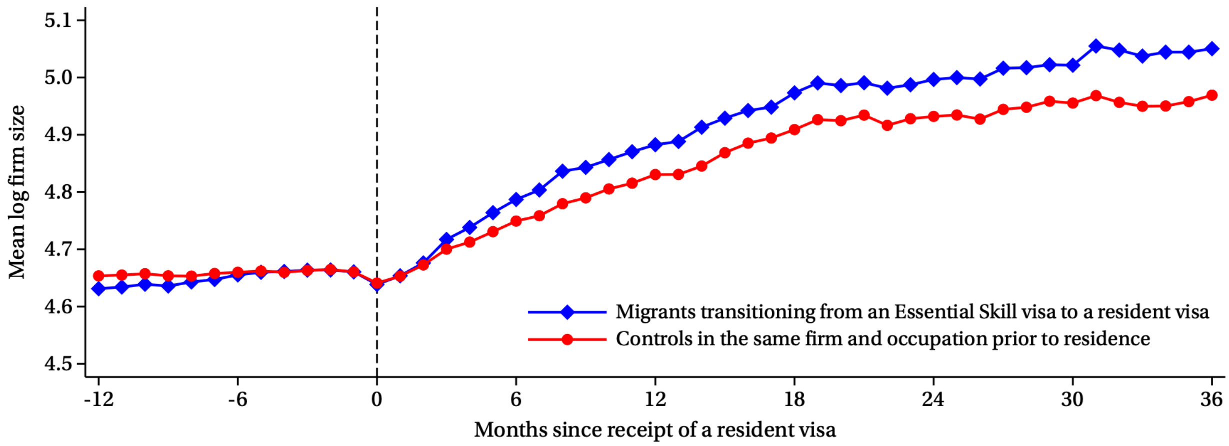


Panel B: Visa status



Panel C: Hours of work

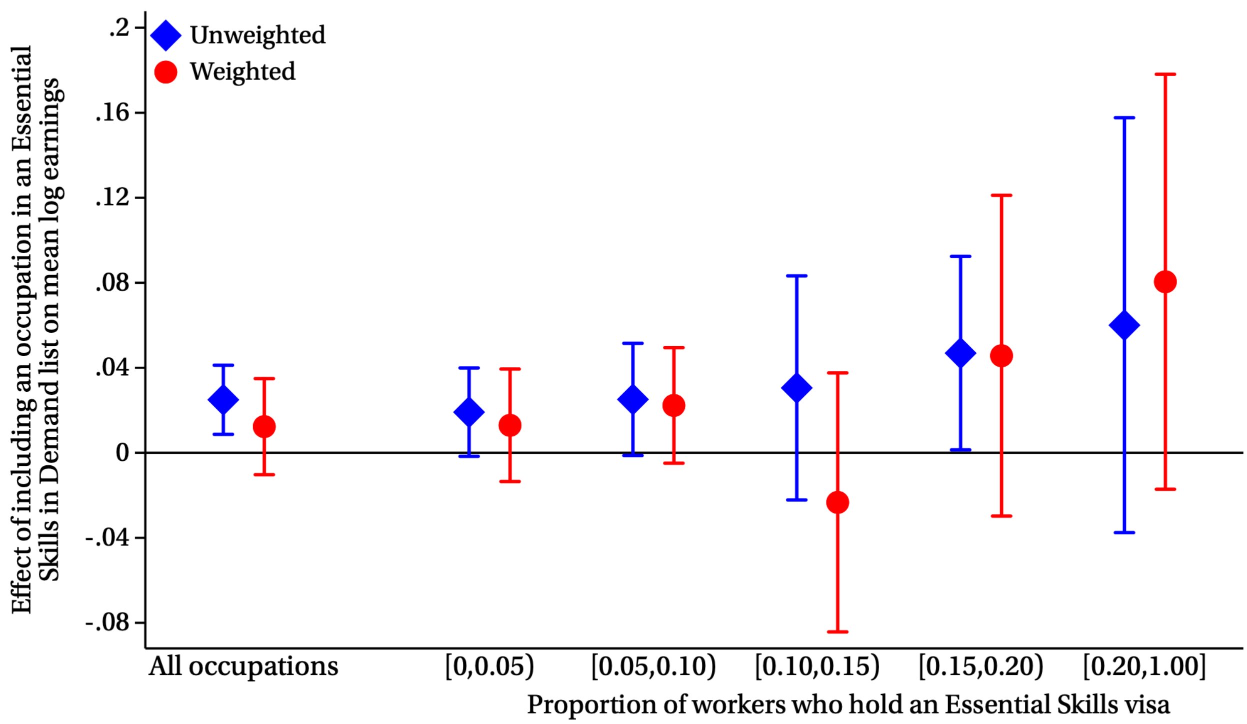
Figure A1 (continued): How employment, visa status, hours of work and firm size change when a migrant receives a resident visa



Panel D: Firm size of work

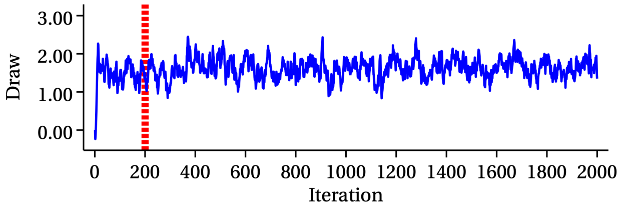
This figure depicts how employment (Panel A), visa status (Panel B), hours of work (Panel C) and firm size (Panel D) change around the time that a migrant transitions from an Essential Skills visa to a resident visa. The dependent variable in Panel A indicates a worker has positive earnings. The dependent variable in Panel B indicates that the worker held either a resident visa or citizenship at the start of the month. The dependent variable in Panel C is a worker's self-reported usual weekly hours of work. The dependent variable in Panel D is the mean log number of employees, where the mean is taken across all firms in which a worker is employed in a given month. The blue series comprises only workers who held an Essential Skills visa for at least 18 months prior to receiving a resident visa. The red series comprises workers who worked in the same firm and occupation as the modal migrant in the month prior to the modal migrant receiving a resident visa. See Section 3 for data sources and Subsection 5.1 for details on the analytic sample.

Figure A2: How including an occupation on an Essential Skills in Demand list affects earnings, by the proportion of that occupation who hold an Essential Skills visa

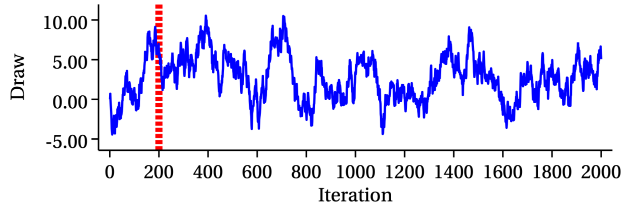


This figure depicts the conditional average effect of including an occupation on an Essential Skills in Demand list, by the proportion of that occupation who hold an Essential Skills visa. The proportion who hold an Essential Skills visa is a time-invariant variable, calculated by taking the mean across all monthly observations for each occupation. Occupation-specific effects are calculated using the estimator presented in Appendix D. In the blue series, these effects are averaged without weighting; in the red series, these effects are weighted by the number of workers in the occupation during the month it was added to or removed from an Essential Skills in Demand list. Bars represent 95% confidence intervals. See Section 6 for details on the analytic sample.

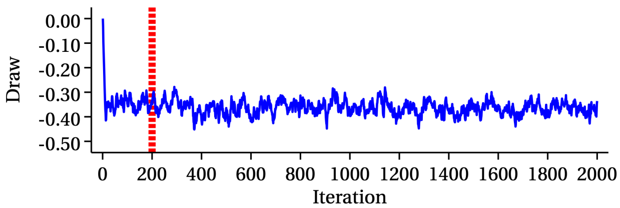
Figure A3: Trace plots for the Stayers Average Treatment Effect



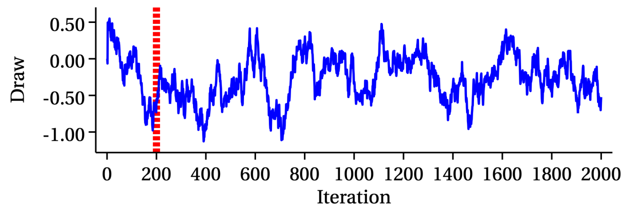
Panel A: β_0



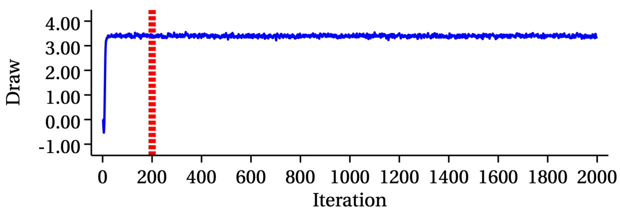
Panel B: β_1



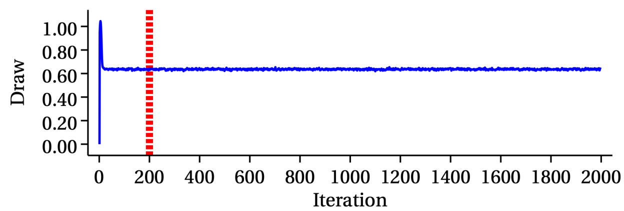
Panel C: β_2



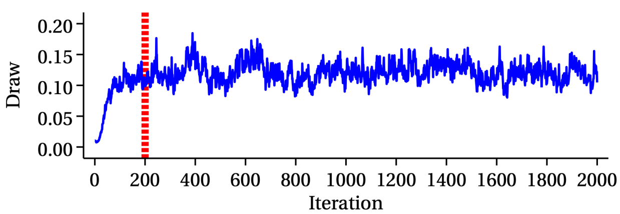
Panel D: β_3



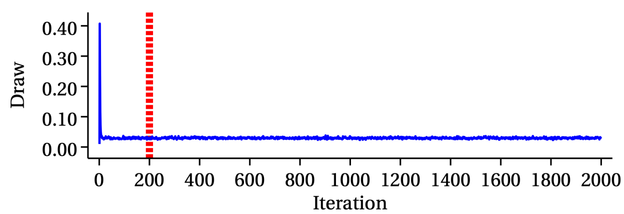
Panel E: α_0



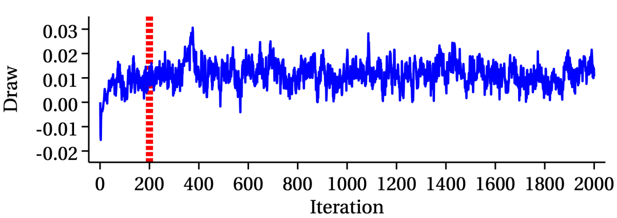
Panel F: α_1



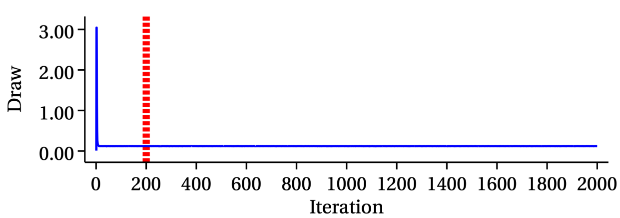
Panel G: $\mathrm{var}(\delta)$



Panel H: $\mathrm{var}(\gamma)$



Panel I: $\mathrm{cov}(\delta,\gamma)$



Panel J: $\mathrm{var}(\epsilon)$

This figure depicts trace plots for the Gibbs sampler used to calculate the posterior distribution of the Stayers Average Treatment Effect. Each panel depicts the path of one hyperparameter. See Appendix C for details on the estimation algorithm. The area left of the red dashed line is the burn-in period, which is discarded when forming the posterior distribution described in tables 6 and A5.

Table A1: Cross-sectional wage regressions

Panel A: 2013 Census; those with 12 months of earnings at a unique employer

	(1)	(2)	(3)	(4)	(5)	(6)
Fixed effects:	None	None	Occupation	Firm	Occ. × firm	Occ. × firm
Holds an Essential Skills visa	-0.194	-0.184	-0.048	-0.025	-0.010	-0.032
	[0.006]	[0.005]	[0.004]	[0.005]	[0.006]	[0.006]
Log hours of work	—	0.974	0.829	0.965	0.641	—
		[0.004]	[0.004]	[0.004]	[0.006]	
Age:						
30-39	—	0.280	0.221	0.239	0.196	—
		[0.001]	[0.001]	[0.001]	[0.002]	
40-49	—	0.385	0.311	0.344	0.282	—
		[0.001]	[0.001]	[0.001]	[0.002]	
50-59	—	0.367	0.310	0.343	0.293	—
		[0.001]	[0.001]	[0.001]	[0.002]	
60+	—	0.329	0.283	0.324	0.277	—
		[0.002]	[0.002]	[0.002]	[0.002]	
Highest qualification:						
High school	—	0.167	0.065	0.100	0.036	—
		[0.001]	[0.001]	[0.001]	[0.002]	
Trade certificate or diploma	—	0.290	0.136	0.200	0.084	—
		[0.001]	[0.001]	[0.001]	[0.002]	
Bachelor's degree	—	0.463	0.227	0.330	0.135	—
		[0.002]	[0.002]	[0.002]	[0.002]	
Postgraduate degree	—	0.590	0.304	0.430	0.192	—
		[0.002]	[0.002]	[0.002]	[0.003]	
R^2	0.001	0.333	0.547	0.609	0.835	0.790
Number of fixed effects	0	0	1005	78219	274665	279885
Number of workers	714915	696870	679533	696870	679533	696648

Panel B: 2013 Census; those with 6 months of earnings at a unique employer

	(1)	(2)	(3)	(4)	(5)	(6)
Fixed effects:	None	None	Occupation	Firm	Occ. × firm	Occ. × firm
Holds an Essential Skills visa	-0.157	-0.156	-0.043	-0.003	-0.015	-0.034
	[0.005]	[0.005]	[0.003]	[0.005]	[0.005]	[0.005]
Log hours of work	—	0.964	0.826	0.952	0.645	—
		[0.004]	[0.004]	[0.004]	[0.005]	
Age:						
30-39	—	0.279	0.216	0.235	0.188	—
		[0.001]	[0.001]	[0.001]	[0.001]	
40-49	—	0.386	0.305	0.337	0.272	—
		[0.001]	[0.001]	[0.001]	[0.002]	
50-59	—	0.368	0.303	0.336	0.281	—
		[0.001]	[0.001]	[0.001]	[0.002]	
60+	—	0.329	0.275	0.313	0.262	—
		[0.002]	[0.001]	[0.002]	[0.002]	
Highest qualification:						
High school	—	0.160	0.061	0.095	0.034	—
		[0.001]	[0.001]	[0.001]	[0.001]	
Trade certificate or diploma	—	0.281	0.131	0.192	0.082	—
		[0.001]	[0.001]	[0.001]	[0.002]	
Bachelor's degree	—	0.457	0.220	0.320	0.132	—
		[0.002]	[0.002]	[0.002]	[0.002]	
Postgraduate degree	—	0.589	0.297	0.420	0.188	—
		[0.002]	[0.002]	[0.002]	[0.003]	
R^2	0.001	0.333	0.542	0.609	0.830	0.785
Number of fixed effects	0	0	1008	88098	314133	320070
Number of workers	833061	811881	792060	811881	792060	812175

Panel C: 2018 Census; those with 12 months of earnings at a unique employer						
	(1)	(2)	(3)	(4)	(5)	(6)
Fixed effects:	None	None	Occupation	Firm	Occ. × firm	Occ. × firm
Holds an Essential Skills visa	-0.222	-0.189	-0.049	-0.048	-0.004	-0.026
	[0.003]	[0.003]	[0.003]	[0.004]	[0.005]	[0.005]
Log hours of work	—	0.980	0.806	0.940	0.623	—
		[0.004]	[0.004]	[0.004]	[0.005]	
Age:						
30-39	—	0.265	0.209	0.232	0.194	—
		[0.001]	[0.001]	[0.001]	[0.002]	
40-49	—	0.379	0.308	0.350	0.288	—
		[0.001]	[0.001]	[0.001]	[0.002]	
50-59	—	0.372	0.314	0.356	0.303	—
		[0.001]	[0.001]	[0.001]	[0.002]	
60+	—	0.323	0.285	0.330	0.287	—
		[0.002]	[0.001]	[0.002]	[0.002]	
Highest qualification:						
High school	—	0.155	0.061	0.095	0.035	—
		[0.001]	[0.001]	[0.001]	[0.002]	
Trade certificate or diploma	—	0.267	0.121	0.185	0.076	—
		[0.002]	[0.001]	[0.001]	[0.002]	
Bachelor's degree	—	0.408	0.196	0.306	0.119	—
		[0.002]	[0.002]	[0.002]	[0.002]	
Postgraduate degree	—	0.545	0.280	0.416	0.176	—
		[0.002]	[0.002]	[0.002]	[0.002]	
R^2	0.003	0.314	0.553	0.592	0.832	0.786
Number of fixed effects	0	0	1017	84003	295482	296244
Number of workers	772842	770367	750789	770367	750789	753039

Panel D: 2018 Census; those with 6 months of earnings at a unique employer						
	(1)	(2)	(3)	(4)	(5)	(6)
Fixed effects:	None	None	Occupation	Firm	Occ. × firm	Occ. × firm
Holds an Essential Skills visa	-0.180	-0.161	-0.039	-0.027	0.002	-0.012
	[0.003]	[0.003]	[0.002]	[0.003]	[0.004]	[0.004]
Log hours of work	—	0.979	0.809	0.933	0.629	—
		[0.004]	[0.004]	[0.004]	[0.005]	
Age:						
30-39	—	0.263	0.202	0.228	0.184	—
		[0.001]	[0.001]	[0.001]	[0.001]	
40-49	—	0.376	0.298	0.343	0.275	—
		[0.001]	[0.001]	[0.001]	[0.001]	
50-59	—	0.369	0.303	0.347	0.289	—
		[0.001]	[0.001]	[0.001]	[0.001]	
60+	—	0.320	0.273	0.319	0.270	—
		[0.001]	[0.001]	[0.001]	[0.002]	
Highest qualification:						
High school	—	0.148	0.058	0.091	0.034	—
		[0.001]	[0.001]	[0.001]	[0.001]	
Trade certificate or diploma	—	0.258	0.117	0.178	0.076	—
		[0.001]	[0.001]	[0.001]	[0.002]	
Bachelor's degree	—	0.400	0.187	0.294	0.114	—
		[0.002]	[0.002]	[0.002]	[0.002]	
Postgraduate degree	—	0.539	0.271	0.403	0.170	—
		[0.002]	[0.002]	[0.002]	[0.002]	
R^2	0.002	0.318	0.546	0.590	0.826	0.779
Number of fixed effects	0	0	1020	94686	341199	342120
Number of workers	909108	906093	883395	906093	883395	886146

Table A1 reports cross-sectional regressions in which the dependent variable is an employee's log earnings across either the 6 months or 12 months prior to either the 2013 Population Census or the 2018 Population Census. These regressions are estimated using full-time paid employees with a unique employer. See Section 3 for data sources and see Section 4 for a detailed description of the sample.

Table A2: How receiving a resident visa affects earnings

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	All occupations	On a skill list	Not on a skill list	Managers	Professionals	Technicians & trades	Community & personal services	Clerical & admin.	Sales	Machinery & drivers	Laborers
Panel A: Dependent variable is log earnings											
Effect of a resident visa β	0.002	-0.001	0.004	0.001	-0.003	-0.005	0.042	0.009	0.007	0.030	0.011
	[0.003]	[0.004]	[0.005]	[0.009]	[0.005]	[0.004]	[0.011]	[0.019]	[0.031]	[0.023]	[0.019]
Coefficient on baseline δ	0.433	0.377	0.496	0.544	0.455	0.339	0.431	0.508	0.545	0.597	0.322
	[0.020]	[0.029]	[0.023]	[0.037]	[0.050]	[0.025]	[0.047]	[0.151]	[0.114]	[0.097]	[0.058]
Number of observations	275550	160176	115377	36399	84495	113193	22650	4731	4128	4389	5568
Number of new residents	11292	6168	5124	1506	3378	4476	1188	219	174	147	201
Number of baseline firms	5769	3465	2907	1197	1605	2757	453	189	123	120	138
Panel B: Dependent variable is colleagues' mean log earnings											
Effect of a resident visa β	0.008	0.004	0.014	0.019	0.001	0.004	0.036	0.036	0.014	0.033	0.008
	[0.002]	[0.002]	[0.003]	[0.006]	[0.003]	[0.003]	[0.007]	[0.010]	[0.023]	[0.014]	[0.012]
Coefficient on baseline δ	0.327	0.282	0.395	0.414	0.346	0.281	0.450	0.559	1.180	-2.271	0.448
	[0.054]	[0.066]	[0.083]	[0.091]	[0.155]	[0.078]	[0.149]	[0.153]	[1.324]	[0.739]	[0.179]
Number of observations	218892	128655	90240	27639	69621	89781	16911	3867	3072	3675	4323
Number of new residents	9555	5397	4158	1287	2874	3891	858	183	156	132	180
Number of baseline firms	5001	3021	2484	1020	1392	2397	366	156	108	108	123

Table A2 reports estimates of the effect of receiving a resident visa using the matching estimator in Equation (2). Column (1) is estimated using all workers, while later columns are estimated on subsets of workers based on their occupation in the month prior to them receiving a resident visa. Columns (2) and (3) distinguish occupations by whether they were on an Essential Skills in Demand list during the month prior to the receipt of the resident visa. The dependent variable in Panel A is log earnings, and the dependent variable in Panel B is the mean log earnings of contemporaneous colleagues (among worker-month observations with a unique employer, excluding spell starts and ends).Standard errors are clustered at the level of the baseline firm. See Section 3 for data sources and see Subsection 5.1 for details on the analytic sample and estimator.

Table A3: Lottery sample summary statistics

	All entrants	Successful entrants	Unsuccessful entrants
Percentage from:			
Fiji	0.684	0.403	0.693
Samoa	0.160	0.290	0.156
Tonga	0.128	0.226	0.125
Tuvalu or Kiribati	0.027	0.079	0.026
Average age	34.8	34.3	34.8
Average number of dependants	1.9	2.0	1.9
Average months in New Zealand	44.5	47.3	44.4
Percentage men	0.791	0.746	0.793
Percentage in Auckland	0.456	0.444	0.456
Percentage urban	0.665	0.660	0.665
In the six months prior to the lottery:			
Percentage employed in NZ	0.953	0.968	0.952
Average monthly NZ earnings	3927.9	3811.8	3931.6
Percentage starting a new job	0.090	0.145	0.088
Percentage with a resident visa in			
1 year	0.062	0.603	0.044
2 years	0.128	0.794	0.107
3 years	0.212	0.852	0.190
Num. entries	6123	189	5934
Num. workers	3465	186	3363
Num. lotteries	48	33	48

Table A3 describes our sample of workers who entered a lottery for either a Pacific Access Category visa or a Samoan Quota visa while already in New Zealand on an Essential Skills visa. Residents of Tuvalu and Kiribati entered separate lotteries; their proportions are combined in this table only to preserve anonymity. The average number of dependants is measured only among primary entrants. Age, number of dependants, months in New Zealand, gender, and location are all measured at the time of the lottery. Earnings are expressed in 2020 New Zealand dollars.

Table A4: How winning a resident visa lottery affects other outcomes

	(1)	(2)	(3)
Dependent variable:	Employed in New Zealand	Located in New Zealand	Colleagues' log earnings
Second-stage slope coefficient	0.027	0.066	-0.011
	[0.037]	[0.021]	[0.034]
First-stage slope coefficient	0.676	0.676	0.714
	[0.027]	[0.027]	[0.027]
First-stage F statistic	635.464	635.464	692.279
Control mean	0.851	0.908	8.305
Number of observations	149439	149439	109746
Number of workers	3465	3465	3150

Table A4 reports two-stage least squares estimates of the linear IV system (5). Standard errors are clustered at the worker level. The system is estimated using months 12–36 following the lottery; Column (3) additionally restricts to observations with positive earnings and a unique firm, and excludes workers' first and last months at a firm. Standard errors are clustered at the worker level. See Section 3 for data sources and Subsection 5.2 for details on the analytic sample and estimator.

Table A5: Estimates of a parametric model of earnings and job-switching

	(1)	(2)
Job-switching depends on:	Current earnings	Current and lagged earnings
Stayers average treatment effect	-0.050	-0.027
	[0.029]	[0.029]
Job-switching selection equation		
β_0	1.624	1.554
	[0.253]	[0.349]
β_1	2.761	-0.585
	[2.723]	[2.544]
β_2	-0.364	-0.060
	[0.027]	[0.070]
β_3	-0.277	-0.226
	[0.295]	[0.491]
β_4	—	-0.299
		[0.049]
β_5	—	0.305
		[0.461]
$\mathrm{var}(\delta)$	0.121	0.112
	[0.016]	[0.015]
Earnings D.G.P.		
α_0	3.394	3.424
	[0.045]	[0.045]
α_1	0.636	0.633
	[0.005]	[0.005]
$\mathrm{var}(\epsilon)$	0.121	0.121
	[0.001]	[0.001]
$\mathrm{var}(\gamma)$	0.030	0.029
	[0.002]	[0.002]
$\mathrm{cov}(\delta, \gamma)$	0.012	0.000
	[0.005]	[0.005]
Number of observations	62361	61704
Number of workers	2982	2982

Table A5 reports the estimated primitives of the parametric model of job-switching and earnings described in Subsection 5.3, with which we infer how winning a resident visa lottery affects the earnings of those who remain at their baseline firms. Posterior standard deviations are included in brackets.

Table A6: Sample of occupation-month observations for our analysis of Essential Skills in Demand lists

	Occupations never on an Essential Skills in Demand list	Occupations sometimes on an Essential Skills in Demand list	Occupations always on an Essential Skills in Demand list
Mean age	40.8	40.7	39.9
Proportion men	0.598	0.665	0.734
Mean real monthly earnings	5218.4	7376.3	8520.1
Proportion with an Essential Skills visa	0.135	0.153	0.152
Proportion with a non-resident and non-Essential Skills visa	0.081	0.072	0.086
Proportion with a resident visa or citizenship	0.784	0.775	0.762
Number of workers	1690026	633693	372288
Number of occupation-month observations	107868	30054	12456
Number of occupations	750	198	81

Table A6 describes the occupation-by-month sample used in Section 6 to analyse the effect of including an occupation on an Essential Skills in Demand list. The table reports unweighted means of occupation-month observations, where the occupation-month observation is itself either a mean or a proportion. These means are estimated using only worker-month observations with nonmissing occupation and thus over-represent Essential Skills migrants, who are more likely to have their occupation listed in administrative data. See Section 3 for data sources and Section 6 for details on the analytic sample.

Table A7: Worker characteristics, by the effect of expanding migrants' job options on their log wages

Panel A: Effects on Migrants				
Effect of expanding migrants' job options on log wages	Number of migrants	Mean baseline log wage	Mean prop. migrant among initial colleagues	Proportion who benefit
< -0.1	1401	3.917	0.270	1.000
-0.10 to -0.08	129	3.764	0.263	0.977
-0.08 to -0.06	162	3.758	0.246	0.981
-0.06 to -0.04	156	3.804	0.251	1.000
-0.04 to -0.02	162	3.773	0.245	1.000
-0.02 to 0.00	192	3.844	0.223	0.922
0.00 to 0.02	1410	3.803	0.172	1.000
0.02 to 0.04	606	3.754	0.301	1.000
0.04 to 0.06	315	3.745	0.256	1.000
0.06 to 0.08	273	3.662	0.269	1.000
0.08 to 0.10	207	3.752	0.280	1.000
> 0.1	2148	3.591	0.386	1.000
All migrants	7161	3.747	0.285	0.997

Panel B: Effects on Residents				
Effect of expanding migrants' job options on log wages	Number of residents	Mean baseline log wage	Mean prop. migrant among initial colleagues	Proportion who benefit
< -0.1	222	4.035	0.003	0.784
-0.10 to -0.08	24	3.437	0.018	0.750
-0.08 to -0.06	36	3.956	0.006	0.750
-0.06 to -0.04	114	3.751	0.000	0.316
-0.04 to -0.02	1161	3.757	0.000	0.023
-0.02 to 0.00	374982	3.899	0.000	0.000
0.00 to 0.02	187593	3.974	0.012	1.000
0.02 to 0.04	6567	3.837	0.156	1.000
0.04 to 0.06	2679	3.789	0.166	0.997
0.06 to 0.08	1893	3.773	0.213	0.997
0.08 to 0.10	783	3.842	0.231	0.992
> 0.1	2535	3.722	0.313	0.981
All residents	578589	3.920	0.009	0.350

Table A7 depicts the heterogeneous effects of allowing migrants to work at any firm, using the equilibrium simulations discussed in Section 8. Each row describes a different set of workers, with sets defined by the worker's visa status and by the effect that allowing migrants to work at any firm would have on the worker's log wage. 'Mean baseline log wage' measures the mean log hourly wage in the initial equilibrium. To calculate 'mean prop. migrant among initial colleagues' we first calculate, for each firm-occupation cell, the migrant proportion of labor mass in the initial equilibrium (treating labor as continuous). We then assign this proportion to each worker on the basis of their initial firm, and take the mean across workers. 'Proportion who benefit' measures the proportion of workers with higher utility in the counterfactual equilibrium than in the initial equilibrium.

Table A8: Firm characteristics, by the effect of expanding migrants' job options on their log profits

Effect of expanding migrants' job options on log profits	Number of firms	Mean employees	Mean productivity	Mean prop. migrant among employees	Prop. which employ any migrants	Mean effect on log mean wages	Mean effect on log employees
< -0.1	858	10.83	4.760	0.284	1.000	0.091	-0.220
-0.10 to -0.08	177	20.76	4.708	0.140	0.983	0.025	-0.080
-0.08 to -0.06	225	26.06	4.766	0.124	1.000	0.017	-0.064
-0.06 to -0.04	294	32.36	4.723	0.115	0.959	0.011	-0.046
-0.04 to -0.02	537	30.66	4.787	0.083	0.966	0.005	-0.027
-0.02 to 0.00	8835	16.06	4.890	0.024	0.800	0.000	-0.001
0.00 to 0.02	30351	16.45	4.797	0.011	0.276	0.000	0.004
0.02 to 0.04	2337	8.80	4.722	0.005	0.135	-0.003	0.029
0.04 to 0.06	870	7.71	4.750	0.003	0.117	-0.004	0.049
0.06 to 0.08	408	5.99	4.743	0.004	0.125	-0.005	0.071
0.08 to 0.10	237	4.18	4.720	0.002	0.076	-0.008	0.091
> 0.1	489	3.97	4.778	0.001	0.031	-0.012	0.163
All firms	45618	15.75	4.808	0.020	0.395	0.001	0.002

Table A8 depicts the heterogeneous effects of allowing migrants to work at any firm, using the equilibrium simulations discussed in Section 8. Each row describes a different set of firms, with sets defined by the effect that allowing migrants to work at any firm would have on the firm's log profits. 'Mean employees' measures the mean number of employees in the initial equilibrium. To calculate 'mean productivity' we first measure firm-level productivity by taking the firm-level mean of the firm-occupation productivity terms $\phi_{o,f}$, weighting each occupation by its mass of employees in the initial equilibrium. We then average these firm-level productivity terms across the relevant set of firms. 'Mean prop. migrant among employees' measures the mean proportion of firms' employees who are migrants in the initial equilibrium. 'Prop. which employ any migrants' measures the proportion of firms which employ any migrants in the initial equilibrium. 'Mean effect on log mean wages' measures the mean firm-specific effect of expanding migrants' job options on log mean wages. 'Mean effect on log employees' measures the mean firm-specific effect of expanding migrants' job options on log employee count.