Volume Title: High-Skilled Migration to the United States and Its Economic Consequences

Volume Author/Editor: Gordon H. Hanson, William R. Kerr, and Sarah Turner, editors

Volume Publisher: University of Chicago Press

Volume ISBNs: 978-0-226-52552-5 (cloth); 978-0-226-52566-2 (e-ISBN)

Volume URL: http://www.nber.org/books/hans-4

Conference Date: July 10-11, 2016

Publication Date: May 2018

Chapter Title: Firm Dynamics and Immigration: The Case of High-Skilled Immigration

Chapter Author(s): Michael E. Waugh

Chapter URL: http://www.nber.org/chapters/c13843

Chapter pages in book: (p. 205 – 238)

6

Firm Dynamics and Immigration
The Case of High-Skilled Immigration

Michael E. Waugh

6.1 Introduction

How does immigration affect relative wages, output, and welfare? How are the gains (or losses) from changes in immigration policy accrued over time? I show how the dynamics of the firm yields new insights into the short- and long-run responses of relative wages, output, and consumption to changes in immigration policy.

The theoretical starting point is a dynamic, heterogeneous firm model, as in Hopenhayn (1992), Hopenhayn and Rogerson (1993), and Melitz (2003). Firms are monopolistic competitors that differ in their productivity and firms must pay a per-period fixed cost of operation. There is free entry, and firms endogenously exit when the value of operating is less than that of exiting. I model labor demand by following the immigration literature: firms employ a constant elasticity of substitution composite of skilled and unskilled labor. The key departure that I entertain is the possibility that the skill intensity of production varies with a firm's productivity.

Nontrivial dynamics in relative wages arise from the interaction between firm productivity and skill. I analytically show how this interaction breaks the "industry standard" constant-elasticity relationship between wages and

skill supply (which is widely used in the immigration literature; see, e.g., Card [2009]; Borjas [2014]). In particular, the deviation from the constant-elasticity benchmark depends on the distribution of firms; thus, the change in relative wages with respect to a change in labor supply depends, in part, on the evolution of the distribution of firms. In contrast, if there is no interaction between firm productivity and skill, then firm heterogeneity and firm dynamics play no role in shaping the aggregate skill premium and its response to immigration.

I quantitatively illustrate these issues by evaluating two types of policies: a "neoliberal" policy that expands the H-1B visa program and a "nationalistic policy" that eliminates it. The H-1B program is a large visa program for the temporary immigration of skilled labor to the United States. At current rates, the general quota allows up to 65,000 employment-based immigrants per year, with an additional 20,000 visas for those with advanced degrees from US universities. Business leaders (e.g., Mark Zuckerberg of Facebook and Bill Gates of Microsoft) argue that this cap is constraining and that an expansion of the H-1B visa program is vital for their firms (and others) to expand, grow, and innovate. However, some policymakers have expressed concerns that this program incentivizes firms to substitute into cheaper, immigrant labor at the cost of displacing domestic workers and/or lowering their wages.

The neoliberal policy that I evaluate is the proposed Immigration Innovation Act of 2015[1] or "I-Squared," which sought to triple the number of H-1B visas per year. To evaluate this policy proposal, I calibrate the parameters of the model to match key properties of firms and labor market outcomes in the US economy. I project forward how the I-Squared Act affects the stock of skilled workers in the United States. I then compute the transition path of the economy in response to an unanticipated adoption of the I-Squared Act.

This reform generates nontrivial, short-run dynamics in relative wages that differ from their long-run dynamics. In the short run, the wage premium of high-skilled to low-skilled workers contracts *more* than would be predicted by a standard, static constant elasticity of substitution (CES) model. The reason is that immigration induces firm entry and that entrants are likely to be low-skill intensive. Thus, entry bids up the relative price of low-skilled labor and the skill premium decreases by more than a static CES model would predict. This process dissipates as entrants become incumbents and the economy converges to its new stationary equilibrium.

The value added of the model is that I can evaluate the level effects on wages, output, consumption, and welfare—not just the distributional effects. I show that the I-Squared Act generates essentially no negative impact on the level of high-skilled wages. Furthermore, this leads to small increases

---

1. See https://www.congress.gov/bill/114th-congress/senate-bill/153.

in aggregate gross domestic product (GDP) in year 1, and a 1.5 percentage point increase in GDP fifteen years out.

These gains arise from both a scale and an aggregate productivity effect from adding more skilled labor. These gains are analogous to the gains from trade emphasized in the monopolistic-competition models of Krugman (1980) or Melitz (2003). The surprising result is the *speed* at which these gains are realized—the no-negative wage impact comes from firms entering quickly. The entry of firms is typically thought of as a long-run effect (see, e.g., di Giovanni, Levchenko, and Ortega [2015], who explain this logic well). In my model, however, this benefit is felt in year 2—not in the "long run." The reason is the dynamic, forward-looking nature of the firm. And this detail—the accrual of long-run benefits today—would be overlooked in a steady state to steady state comparison.[2]

The flip side of all these "good" outcomes—higher wages and higher output—is that they come from firm entry, and firm entry must be paid for. There is investment today in the creation of firms to prepare for a larger labor force in the future—and this investment comes at the cost of consumption. Under an assumption about the distribution of profits across workers, the I-Squared Act leads to a drop in consumption of a 0.5 percentage point for both workers (in year 1) and stays depressed relative to previous levels for at least four years. This experiment makes an important conceptual point about who bears the burden of the adjustment to the I-Squared policy—it's the owners of the firm, not the workers.

The nationalistic policy that I evaluate is a complete elimination of the H-1B visa program.[3] Mimicking the results above, this policy delivers the following: the skill premium expands with negative effects on low-skilled workers; firms' exit and entry contracts; output contracts, and yet consumption overshoots, as there is a reduction in investment in new firms.

A unique outcome of the nationalistic policy is its unintended negative effects on the wages of low-skilled workers in the short run. As with the I-Squared policy, these consequences work through a change in the distribution of firms. Due to the elimination of the H-1B program, firms foresee a smaller market that results in less entry and more exit. Since entrants and exiting firms are less productive and low-skill intensive, low-skilled workers' wages contract as the demand for their labor services erodes.

As with the I-Squared policy, the normative implications of the nationalistic policy are subtle. While output declines through the scale and productivity effect, consumption increases in the short run. The issue here is that lower wages and lower output come from less entry and exit of firms. This is

2. Lee (2016) also focuses on the transition of the economy in response to changes in immigration. However, the dynamics of the economy arise from workers' life-cycle motives.

3. One current policy proposal is the High-Skilled Integrity and Fairness Act of 2017 (https://www.congress.gov/bill/115th-congress/house-bill/670/text), which seeks to increase the minimum salary requirement for H-1B visa holders.

because there is less need for investment in new firms, as the economy has too many firms given the shrinkage of the labor force (today and in the future). And the reduction in investment comes at the benefit of higher profits and consumption in the short run. While this effect mitigates the negative consequences of a nationalistic policy, it does highlight the following point: the negative consequences of a nationalistic immigration policy are borne by the workers—not by the owners of the firms.

This chapter provides an answer to some fundamental and unanswered questions: What are the distributional and aggregate effects of immigration? Regarding the distributional effects, there appears to be a wide range of answers within the literature. Estimating the distributional effects relies upon estimates of the elasticity of substitution between workers' types. These estimates seem to give wide-ranging answers depending upon the source of identifying variation, categorization of worker types, the instrument, and so forth (see, e.g., the discussions in Card [2009] or Borjas [2014]). Some estimates suggest near-zero impacts on relative wages, and some are larger.

One explanation for this discrepancy is that there are non-labor-market adjustments taking place in the background (see, e.g., the discussion in Lewis [2013]). I contribute to this line of thought by exploring one margin of non-labor-market adjustment: how immigration affects firms' entry and exit decisions. A key result is that changes in labor supply shift labor demand and lead to different short- and long-run wage responses as firms enter and use different skill mixes relative to incumbents.

Regarding the aggregate effects, the typical approach in the immigration literature is to treat the relative wage response (given an estimated elasticity of substitution) as a sufficient statistic for the outcome from immigration. Under certain restrictions on technologies, this is appropriate. However, in my model, the dynamics of the firm lead to outcomes in which the welfare effects of immigration are not captured by changes in relative wages. As discussed above, the adjustment of firms leads to substantial changes in consumption in the short run, even if the wage effects from immigration are negligible.

This chapter owes a large, intellectual debt to the trade literature and its emphasis on the role of the firm. The work of Bernard and Jensen (1999), Melitz (2003), and Bernard et al. (2003) very much focused the trade literature on the role of firms and their adjustments in understanding the positive and normative implications of trade.

Specifically, this chapter builds on two ideas discussed in the recent trade literature. First, my model shares the "skill-biased productivity mechanism" emphasized in Burstein and Vogel (forthcoming), with the key difference that I study the dynamic effects of a supply shock (immigration) rather than on a demand shock (opening to trade). This chapter also borrows from the idea that firm dynamics lead to horizon-varying trade elasticities, as in the work of Ruhl (2008) and Alessandria, Choi, and Ruhl (2014). In the immi-

gration context, I show when firm heterogeneity matters (and does not) and how the characteristics of entering firms affect the elasticity across worker types over different time horizons.

6.2   Model

I outline the model below by describing the consumers (who are also the workers) and the firms. The interesting economics lie with the firms—specifically, how skill mix varies with firm type and the dynamic choices of the firm.

6.2.1   Time and Consumers

Time is discrete and evolves for the infinite horizon. Consumers have the following preferences:

$$(1) \qquad U = \sum_{t=0}^{\infty} \beta c_t,$$

where U is the present discounted value of the instantaneous utility of consuming the final consumption good, and $\beta \in (0,1)$ is the discount factor. The final consumption good is an aggregate bundle of varieties, aggregated with a CES function:

$$(2) \qquad c_t = \left[\int_{M(t)} c_t(\omega)^{(\sigma-1)/\sigma} \, d\omega \right]^{\sigma/(\sigma-1)},$$

where $c_t(\omega)$ is consumption of individual variety ω. The parameter σ controls the elasticity of substitution across variety. The measure M defines the endogenous set of varieties consumed.

I abstract from any decisions of consumers to hold or accumulate assets. Consumers simply consume given their income in each period. Since consumers are the workers and the owners of the firm, their income available for consumption comes from both labor earnings and profits from firms.

6.2.2   Firms

There is a continuum of firms that are heterogeneous in productivity, that are monopolistic competitors on the product market, and that face competitive labor markets.[4] Dropping the time index for clarity, firms producing individual varieties have technologies

$$(3) \qquad q(\omega) = z \left[\phi_s(z) \ell_s^{(\theta-1)/\theta} + \phi_u \ell_u^{(\theta-1)/\theta} \right]^{\theta/(\theta-1)},$$

where z is a firm's productivity, ℓ_s and ℓ_u are skilled and unskilled labor, the ϕs are the skill weights, and θ is the elasticity of substitution between labor types.

The production technology in equation (3) is similar to the aggregate,

4. Competitive labor markets are easy to work with. However, in the context of the H-1B program, this assumption abstracts from important details of the labor market for H-1B visa holders. In particular, that the H-1B program ties workers to firms for the duration of the visa.

"nested CES" structure of different skill types used in the immigration literature (see, e.g., Card 2009; Borjas 2014). The key difference is that skill intensity—the ϕs—may vary with firm productivity. For example, if $\phi_{s}(z) > 0$, then skilled workers are relatively more productive in high-productivity firms, leading to a complementarity between skill and productivity across firms. This possibility is discussed in more depth below. This specification is similar to the production function in Burstein and Vogel's (forthcoming) study of the skill premium and international trade.

Consumer preferences in equation (2) imply that a firm producing variety ω faces the following demand curve:

$$(4) \qquad p(\omega)^{-\sigma}\left(\frac{Y}{P^{1-\sigma}}\right),$$

where $p(\omega)$ is the price of the variety, Y is aggregate income (both labor and profits), and P equals the CES price index.

Firms' Choice of Skill Mix. Given the production function in equation (3), a firm's relative demand for skilled and unskilled labor is

$$(5) \qquad \frac{\ell_{s}(z)}{\ell_{u}(z)} = \left(\frac{\phi_{s}(z)}{\phi_{u}}\right)^{\theta}\left(\frac{w_{s}}{w_{u}}\right)^{-\theta},$$

where w_{s} and w_{u} are the competitively determined wages for skilled and unskilled workers. With one exception, this demand curve is relatively standard: relative demand for labor is inversely related to the relative wage with elasticity θ. The exception is that if the ϕs vary with skill level, a firm's relative demand for skill varies with productivity. The demand curve in skill (5) implies that the within-firm shares of high- and low-skilled workers are

$$(6) \qquad \pi_{s}(z) = \frac{\phi_{s}(z)^{\theta}w_{s}^{-\theta}}{\phi_{s}(z)^{\theta}w_{s}^{-\theta} + \phi_{u}^{\theta}w_{u}^{-\theta}} \quad \text{and} \quad \pi_{u}(z) = \frac{\phi_{u}^{\theta}w_{u}^{-\theta}}{\phi_{s}(z)^{\theta}w_{s}^{-\theta} + \phi_{u}^{\theta}w_{u}^{-\theta}}.$$

These share formulas tell us the following: if skilled wages are relatively higher, then firms will employ relatively fewer high-skilled workers. If $\phi_{s}(z) > 0$, then more-productive firms will employ relatively more high-skilled workers than less-productive firms will. And if the ϕs do not vary with skill type, then all firms will employ the same shares of high- and low-skilled workers. Finally, it will be useful to define an index of "skill":

$$(7) \qquad \Phi(z) = \left[\phi_{s}(z)\pi_{s}(z)^{(\theta-1)/\theta} + \phi_{u}\pi_{u}(z)^{(\theta-1)/\theta}\right]^{\theta/(\theta-1)},$$

which is a CES aggregate of the share of different skill types. This is a summary statistic of the skill mix of the workers in a firm with productivity z. If high-productivity firms employ relatively more high-skilled workers, then $\Phi(z)$ will be increasing with the productivity of the firm.

Firms' Choice of Price and Quantity. Given the optimal skill mix, I express a firm's (static) profit-maximization problem as

(8)
$$\max_{p(\omega),\ell} p(\omega)z\Phi(z)\ell - \left(w_u\pi_u(z) + w_s\pi_s(z)\right)\ell.$$

That is, choose an output price and labor units (i.e., number of bodies) to maximize period profits. Period profits are revenues minus the skill-share-weighted costs of employing ℓ labor units. This problem leads to the following optimal price:

(9)
$$p(z) = \frac{\sigma}{\sigma-1}\frac{\left(w_u\pi_u(z) + w_s\pi_s(z)\right)}{z\Phi(z)},$$

where prices are a constant markup over the marginal cost of employing an efficiency unit of labor.[5] Marginal cost is a share-weighted wage bill relative to the firm's productivity, adjusted by the skill mix of the workers employed. Demand for labor units is

(10)
$$\ell(z) = \frac{1}{z\Phi(z)}\left(\frac{\sigma}{\sigma-1}\frac{\left(w_u\pi_u(z) + w_s\pi_s(z)\right)}{z\Phi(z)}\right)^{-\sigma}\left(\frac{Y}{P^{1-\sigma}}\right).$$

The firm's static profit function is

(11)
$$\pi(z) = p(z)z\Phi(z)\ell(z) - \left[w_u\pi_u(z) + w_s\pi_s(z)\right]\ell(z),$$

which I use in the discussion of the firm's dynamic problem below.

6.2.3 Firm Dynamics

Firm-level productivity, z, evolves stochastically according to a N-state Markov chain with transition matrix $\mathcal{P}$ and an associated invariant distribution $\tilde{\mathcal{P}}$. This stochastic process is meant to capture the observed changes in firms' size and profitability over time. Apple started out as a two-man operation, hand-building wooden computers in Silicon Valley; only a decade ago, Nokia and BlackBerry were world leaders in the design and production of mobile phones. In an exogenous manner, this stochastic process mimics these changes in firm size and productivity over time that are seen in the data.

This process implies that in any period there is a measure $\mu(z)$ over productivity types. This measure will partially reflect the stochastic process in $\mathcal{P}$. It is also determined endogenously by the exit and entry decisions of firms. Thus, the distribution of firms over productivity is an equilibrium object and an endogenous outcome of the model.

Exit comes about as firms face a per-period, fixed cost of operation κ, which is denominated in units of the final good. The timing is such that if a

---

5. Note that the interaction between productivity and skill will generate dispersion in revenue-based productivity (or TFPR in the language of Hsieh and Klenow [2009]), that is, $p(z)\Phi(z)z$ is not independent of z. Furthermore, consistent with Foster, Haltiwanger, and Syverson (2008), revenue-based productivity in my model is correlated with physical-based productivity.

firm pays the fixed cost, it operates in the next period. If the firm does not pay this fixed cost, then it operates in this period and then exits.

Entry takes place via a large pool of nonactive firms that may enter the economy by paying an entry cost $P\kappa^e$ to gain an initial productivity draw. After receiving their productivity draw, entering firms are exactly like incumbents. Entrants receive their productivity draw from density $\mathcal{P}_e$.

Given this environment, I discuss an incumbent firm's problem and the value of entry.

Incumbents' Dynamic Problem. Given the static profit functions (and focusing on a stationary equilibrium motion where aggregate state variables are not changing), the problem of an incumbent firm is to choose between continuing to operate next period and exiting. Since firms are owned by consumers, firms choose exit policies to maximize the expected present discounted value of real profits, discounting with interest rate $r = (1/\beta) - 1$. The value function of an incumbent firm is

$$(12) \qquad v(z_i) = \max\left[\pi(z_i) - \kappa + \beta\sum_{j=1}^{N}\mathcal{P}(i, j)v(z_j), \ \pi(z_i)\right],$$

where the value of the firm is the maximum over two objects. The first objects are the static profit minus the fixed operating costs plus the expected, discounted continuation value of the firm. The second object is the static profit of the firm if it exits.

Entrants. The entry protocol implies that the value of entry is

$$(13) \qquad v_e = \sum_{j=1}^{N}\mathcal{P}_e(j)v(z_j) - \kappa_e,$$

where $v(z_j)$ is the value of a firm in equation (12), and $\mathcal{P}_e(j)$ is the probability of a firm receiving productivity level z_j. Thus, this says that the value of entry equals the expected value of operating in the market net of entry costs.

6.2.4   Equilibrium

Given the environment described above, I formally define a stationary equilibrium:

DEFINITION 1. *A stationary equilibrium is a collection of allocations for consumers c; allocations, prices, and exit decisions for firms; allocations of workers across firms; wages $\{w\}_{s,u}$, a mass of entrants M_e, and a measure of incumbents μ, such that*

- consumers', firms', and workers' problem is solved;
- labor demand equals labor supply, for each skill type;
- the measure over incumbents is stationary; and
- the free-entry condition is satisfied.

Essentially, firms and consumers optimize, markets clear, and the economy is stationary. The economy being stationary means that aggregate outcomes and the measure of firms over individual states are constant, but

that individual firms will dynamically move through the productivity distribution, exit, or enter. In the quantitative section, I will study a nonstationary economy as it transits between two stationary equilibria.

6.3   The Aggregate Skill Premium

In this section, I derive the aggregate skill premium and the aggregate elasticity of relative wages to relative supply of skill. This relationship is important because within the aggregate, nested CES structure, it provides the foundation for evaluating and interpreting the distributional effects from immigration. In particular, I show (a) the importance of the complementarily between firm productivity and skill, and (b) the role of firm heterogeneity and dynamics.

To derive the aggregate skill premium and its relationship to aggregate skill supply, I start from the aggregate resource constraint:

(14) $$\sum_i \mu(z_i)\pi_s(z_i)\ell(z_i) = L_s \text{ and } \sum_i \mu(z_i)\pi_u(z_i)\ell(z_i) = L_u.$$

Here, $\pi_s(z_i)$ and $\pi_u(z_i)$ are the within-firm shares of skilled and unskilled labor in equation (6), $\mu(z_i)$ is the measure of firms with productivity type z_i, and $\ell(z_i)$ is the quantity demanded of labor units by firms with productivity z_i. Finally, L_s and L_u are the aggregate supplies of skilled and unskilled labor. All equation (14) says is that firm demand equals aggregate labor supply.

Substitution of equation (6) into the aggregate resource constraint (14) connects the aggregate skill premium and aggregate skill supply. Proposition 1 summarizes the result.

PROPOSITION 1 (THE AGGREGATE SKILL PREMIUM). *Log relative wages relate to aggregate, log relative skill supplies*

(15) $$\log(w_s) - \log(w_u) = \Theta(w_s, w_u, \mu, \ell) - \frac{1}{\theta}\Big[\log(L_s) - \log(L_u)\Big],$$

where

(16)
$$\Theta(w_s, w_u, \mu) = -\frac{1}{\theta}\log\left\{\sum_i \frac{\phi_s(z_i)^\theta \, \mu(z_i)\ell(z_i)}{\phi_s(z_i)^\theta w_s^{-\theta} + \phi_u^\theta w_u^{-\theta}}\right\}$$
$$+ \frac{1}{\theta}\log\left\{\sum_i \frac{\phi_u^\theta \, \mu(z_i)\ell(z_i)}{\phi_s(z_i)^\theta w_s^{-\theta} + \phi_u^\theta w_u^{-\theta}}\right\}.$$

Furthermore, the change in the skill premium with respect to a change in relative skill supply is

(17) $$d\log(w_s) - d\log(w_u) = d\Theta - \frac{1}{\theta}\Big[d\log(L_s) - d\log(L_u)\Big].$$

Proposition 1 yields three important observations. First, the relationship in equation (15) is very similar to the theoretical relationship used in the

immigration literature. Changes in relative labor supply lead to changes in relative wages that connect directly with the elasticity of substitution between labor types. The key difference is that this is not a constant-elasticity relationship. In general, the intercept term $\Theta(w_s, w_u, \mu)$ will vary with the skill supply.[6] A change in the $\Theta(w_s, w_u, \mu)$ term represents a shift in the labor demand curve due to a change in relative skill supply.

Second, Proposition 1 shows why the labor demand curve will shift—it is because of the complementarity between skill and productivity. The easiest way to see this point is to "turn off" the complementarity with ϕ_s independent of z. In this case, the intercept term (16) becomes

$$(18) \quad \Theta = -\frac{1}{\theta}\log\left\{\sum_i \frac{\phi_s^\theta\, \mu(z_i)\ell(z_i)}{\phi_s^\theta w_s^{-\theta} + \phi_u^\theta w_u^{-\theta}}\right\} + \frac{1}{\theta}\log\left\{\sum_i \frac{\phi_u^\theta\, \mu(z_i)\ell(z_i)}{\phi_s^\theta w_s^{-\theta} + \phi_u^\theta w_u^{-\theta}}\right\}.$$

And then, after canceling terms in equation (18), we have

$$(19) \quad \Theta = -\frac{1}{\theta}\log(\phi_s) + \frac{1}{\theta}\log(\phi_u),$$

with all endogenous variables dropping out of the intercept. When there is no complementarity between skill and productivity, the elasticity of relative wages is constant with elasticity $1/\theta$.

The intuition for why complementarity matters is that firms are differentially substituting in to or out of labor types. Thus, the distribution of firms and their labor demands matter. When there is no complementarity, all firms substitute in the exact same way, and, thus, the distribution of firms and their labor demand plays no role.

This latter point is closely related to the "skill-biased productivity mechanism" emphasized in Burstein and Vogel (forthcoming). That is, opening to trade reallocates labor demand from low-productivity, low-skill-intensity firms to high-productivity, high-skill-intensity firms, and this mechanism leads to an increase in the skill premium. Their insight shows up in the intercept term in equation (16): shifts in the *distribution* of labor demand (in their case, caused by trade; in my case, immigration) change the skill premium as long as there is complementarity between productivity and skill.

Third, Proposition 1 says that firm dynamics matter for the dynamics of relative wages only when there is complementarity between skill and productivity. Again, equation (19) shows that the distribution of firms and its evolution "separate" from the change in wages. Thus, to have different short- and long-run wage elasticities, it is necessary to have an interaction between skill and productivity.

6. A useful exercise would be to abstract from dynamics and assume a distribution over the zs—that is, Pareto, as done in the trade literature (see, e.g., Chaney 2008). With the right function form for the $\phi_s(z)$, some insight may be possible. In particular, I conjecture that the intercept term and how it would respond would depend on the Pareto-shape parameter; thus, the variation in firm-level productivity dispersion would modulate the wage response.

Finally, these observations have a very close relationship to the work on capital-skill complementarity and immigration in Lewis (2011, 2013) and, more generally, Krusell et al. (2000). Capital-skill complementarity gives rise to a non-constant-elasticity relationship between relative wages and relative skill in a very similar way to equation (17). The difference here—and the empirical content—is that $d\theta$ term in equation (17) relates to firms and their differential adjustment to the change in labor supply.

6.4   Quantification

This section discusses the calibration of the model, which proceeds in three steps. First, I describe functional form assumptions. I then describe how the parameter values are chosen such that the model can replicate key features of firm dynamics in the data. Finally, I discuss how labor supply evolves in the model and how I implement the I-Squared policy.

6.4.1   Specification of Shock Process and Skill Bias

To completely specify the model, I must take a stand on the nature of the shock process, the initial productivity of entrants, and a specification relating productivity to the complementarity between skill and productivity.

I construct a Markov process over the zs so that in logs, z mimics an AR(1) process with normally distributed innovations. I achieve this via Tauchen's (1986) method. This implies that there are two parameters to calibrate: the autocorrelation parameter, ρ, and the standard deviation of the shocks, σ_z.

The entrants' productivity distribution is a mean shift of the invariant distribution associated with the Markov process described above. Specifically, μ_e will be the mean of log productivity for entrants. If μ_e is a negative number, then entering firms will be less productive (on average) than incumbents.

I parameterize the ϕs in the following way. First, I normalize ϕ_u equally to one. I then assume that $\phi_s(z)$ is a log-linear function of z with intercept α and elasticity γ. This functional form has the feature that if $\gamma > 0$, then high-productivity firms employ a larger share of high-skilled workers relative to low-productivity firms. This functional form closely resembles the specification in Burstein and Vogel (forthcoming).

6.4.2   Calibration of Parameters

The parameters of the model are grouped into two categories. One set of parameters consists of those that are chosen outside of the model. I call these "predetermined parameters." The second set consists of those chosen match model moments with data moments, that is, "calibrated parameters." The latter are chosen to mimic key properties of firms in the cross section and over time.

Predetermined Parameters. The time period in the model is a year. Thus, I set the discount factor, β, to 0.98. This corresponds with an annualized risk-

free real interest rate of 2 percent, which is consistent with recent experience in the US economy.

The value for the demand elasticity, σ, is set to 4. The trade literature has put much effort into estimating this parameter, and the value 4 lies within the middle of the range of recent estimates. The estimates that I prefer come from Simonovska and Waugh (2014a, 2014b). At the lower end of the range are the estimates from Broda and Weinstein (2006), who find that the median elasticity across product categories is around 3. At the upper end of the range are aggregate estimates from Parro (2013) and Caliendo and Parro (2015); using aggregate tariff and trade-flow data, they find values near 5 (see Simonovska and Waugh [2014b] for a discussion of these estimates).

I set the elasticity of substitution across skill types to 3. This parameter is not uncontroversial. Card (2009) reports that estimates of the θ between college and high school workers range from about 2.5 to 4. Ottaviano and Peri (2012) find estimate values of θ that lie between 1.5 and 3. Borjas (2003) estimates an inverse elasticity of around 1.4. Setting θ to 3 is near the upper-middle part of this range.

There is an important caveat regarding the discussion of the elasticity of substitution across skill types. Proposition 1 makes the point that a structural interpretation of these empirical estimates is not clear, as labor demand will shift with changes in labor supply. Thus, the mapping from these estimates to the θ parameter in my model is not obvious. One rationale for picking a value near the upper-middle part of the range is that these elasticities are biased downward in my model (see, e.g., figure 6.2).

The autocorrelation is chosen to match the autocorrelation of establishment size observed in the Synthetic Longitudinal Business Database (LBD) (US Census Bureau 2011). Predetermining this parameter outside the calibration routine simplifies computational matters, with no loss in the model's ability to correctly mimic the persistence seen in the data.

Finally, the entry cost is normalized to 1. The top panel of table 6.1 summarizes the predetermined parameters.

Calibrated Parameters. There are five remaining parameters to calibrate: the standard deviation of the shocks to productivity, the fixed cost of operation, the shift in the entrant distribution, and the intercept and slope for the skill-bias function.

I calibrate these five parameters to match five moments. The first moment is about the size distribution. The Statistics of US Businesses from the US Census Bureau reports data that include firms binned by size with data on the number of firms, the number of establishments, employment, and the annual payroll for most US business. Half of all employment is in firms with more than 500 employees; I abuse terminology here, but I will call this the median firm. The average firm size is about twenty employees. Thus, I target a ratio of the median to mean size of twenty-five. The parameter that is most directly informative about this moment is the standard deviation of the zs.

Table 6.1   **Calibration summary**

Parameter	Value	Source or target
Predetermined parameters		
Discount rate, β	0.98	—
Demand elasticity σ	4.0	—
Skill elasticity θ	3.0	—
Autocorrelation of log z	0.90	Autocorrelation of size, synthetic LBD
Entry cost, κ_e	1.0	Normalization
Calibrated parameters		
Standard deviation of log z	0.20	Ratio of median size to mean ≈ 25
Fixed cost of operation, κ	0.14	Entry rate of 10 percent
Shift in entry distribution, μ_e	−0.13	Probability of survival of entrants after 5 years, 0.50
Intercept of skill-bias function, α	−0.55	Skill premium, 1.90
Slope of skill-bias function, γ	1.00	Size-wage premium, 1.30

The second and third moments are computed using the Synthetic Longitudinal Business Database (US Census Bureau 2011). The entry rate is computed as the new establishments relative to the total number of establishments. This number is computed to be about 10 percent in the later time periods of the data set. Here, I am just focusing on recent experience in the US economy and abstract from the long-run declines in start-up activity as Decker et al. (2014), Hathaway and Litan (2014), and others document.

The survival rate is computed as establishments staring in a given-year period that remain open (over some time horizon) relative to all establishments starting in that year. This number is about 50 percent at a five-year horizon. The parameters most informative about these moments are fixed operating cost, κ, and the shift in the entrant distribution μ_e.

The fourth and fifth moments are the aggregate skill premium and the firm-size-wage premium. The former is computed as the relative earnings of skilled to unskilled workers using the Current Population Survey. Specifically, I compare the median usual weekly earnings of workers with a bachelor's degree or above with those workers with less than a bachelor's degree. This provides the estimate that skilled workers earn 1.89 times that of unskilled workers.

The size-wage premium is determined as follows. Using the Statistics of US Businesses, I compute the payroll divided by employment for those firms with more than 500 employees—I call this the average wage above the median. Then, I compare this to the average wage or workers in firms below 500 employees. For the period from 2010 to 2013, this value is 1.30. That is, the average wage in firms with more than 500 employees is 30 percent larger than in firms with fewer than 500 employees.

The size-wage premium moment speaks directly to the slope of the skill-bias function, γ. In the model, since size and productivity correspond with each other, there must be some skill bias to match the fact that larger firms

pay higher wages. Thus, the calibration finds that high-productivity firms demand and use relatively more high-skilled workers. Consistent with my findings, Burstein and Vogel (forthcoming) find a value of γ near 1 when calibrated to match the skill intensity of Mexican firms.

The bottom panel of table 6.1 summarizes the results.

6.4.3   Labor Supply and Its Dynamics

To compute the initial stationary equilibrium, I use labor endowments from aggregate data. I take L_u to stand for the US labor force with less than a college degree. This value is normalized. I then take L_s to stand for the US labor force with a college degree or higher. This value is set at 57 percent of L_u as seen in recent US data.

I want to use the model to evaluate two different policy proposals. The first policy focuses on the Immigration Innovation Act of 2015 or "I-Squared," which seeks to triple the number of H-1B visas. Current policy in the United States allows for a maximum of 65,000 H-1B visas, with an additional 20,000 visas for foreign graduates of US universities with advanced degrees. Thus, current policy allows up to 85,000 visas per year. The I-Squared Act raises the cap to 195,000 visas per year and eliminates the advanced degree exception. Furthermore, the policy proposal contains "escalators" that restrict the visa increase by 20,000 visas per year until reaching the cap of 195,000.

The second policy proposal is a "nationalistic" policy that restricts the movement of labor into the United States. I model this policy as a complete elimination of the H-1B visa program, allowing existing H-1B visas holders to remain until the expiration of their visa, but preventing H-1B visa holders from transitioning to permanent status.

There are several challenges to evaluating the effects of these policies. First, I need to know about the current stock of H-1B visa holders. Second, I need an estimate of how changes in the flow of immigrants affect the stock over time.[7] Unfortunately, little is known about the current stock of H-1B visa holders and how they transition to permanent status or exit the United States as their visa expires (or before). Thus, to construct an estimate of the current and future stock of H-1B visa holders, I build on the work of Lowell (2000) and make some educated guesses.

I start from the fact that the visa cap has been binding in recent years. The H-1B visa is a three-year visa with an option for an additional three-year extension. Thus, I assume that H-1B visa holders stay the maximum period of six years. At current caps, this implies that the stock of H-1B visa holders

---

7. In a static model or steady state-to-steady state comparison, it may be reasonable to assume that the change in the steady state stock is proportional to the change in the flow. First, one needs to know the original stock to evaluate the level of the effects from this policy. Furthermore, I want to evaluate the transition; thus, I need an estimate of how the stock transitions to the new steady state.

is 510,000. This is consistent with projections of the stock of H-1B visa holders by Lowell (2000) and updated projections by Kerr and Lincoln (2010).

To compute the change in the stock, I need to know how H-1B visa holders may (or may not) transition to permanent status. Lowell (2000) suggests that up to 50 percent of expiring H-1B visa holders transition to permanent status. This may be an exaggeration, as there are numerical caps on those with permanent status. Also, processing time is a nontrivial barrier. This assumption implies that, at current rates, each year 42,500 H-1B visa holders transition to permanent status, while the remaining half exits.

The final issue is to connect the H-1B visa holders that transition to permanent status with the permanent stock of high-skilled workers in the United States. To do so, I assume that the stock of labor evolves according to a simple "perpetual inventory" law of motion, and I infer the rate at which high-skilled workers exit the labor force under the assumption that the stock of high-skilled workers is stationary. Specifically, the stock of high-skilled labor evolves according to

(20) $\qquad L_{s,t+1}^{p} = (1 - \delta)L_{s,t}^{p} + \text{new graduates}_t + \text{H-1B transitions}_t,$

where $L_{s,t}^{p}$ is the stock of permanent, high-skilled workers. This law of motion implies that in the steady state,

(21) $\qquad L_{s,ss}^{p} = \dfrac{1}{\delta}(\text{new graduates}_t + \text{H-1B transitions}_t).$

We know that the current stock of high-skilled workers (net of H-1B visa holders) is about 48.5 million (averaged over 2010–2015). The flow of new graduates entering the workforce is about 1.10 million over the same time period (see, e.g., Spreen 2013). The (guesstimated) flow of H-1B visa holders into the permanent workforce is 42,500. This implies a δ of 2.36 percent.

The total stock of the skilled labor force is

(22) $\qquad L_{s,t} = L_{s,t}^{p} + \text{ stock of H-1B Visas}_t,$

or the sum of permanent residents and the stock of H-1B visa holders at that time.

I use these assumptions to project the effects of immigration policy on labor supply. Two final comments on this procedure are warranted. In equations (20) and (22), I assume that native and foreign-born workers are the same. If high-skilled immigrants are positively selected (as the evidence in Grogger and Hanson [2011] suggests), then this implies that I am missing an additional margin. Specifically, that the number of new effective labor units associated with an increase in immigration is larger than I am estimating. Second, I assume that the skill choice is not responding to shifts in labor supply; Bound, Khanna, and Morales (chapter 4, this volume) evaluate this margin of adjustment within the context of the H-1B visa program.

6.5   I-Squared Policy

This section evaluates the economic effects of the I-Squared policy. Below, I first discuss the effects on relative wages and then the aggregate, level effects, and I conclude with a discussion of welfare.

To compute the effects of this policy, I treat the change in policy as unanticipated from the perspective of firms. After the policy is announced, firms understand what the entire projected path of the work force in figure 6.1, panel B, will be. I then compute the transition path of the economy to its new stationary equilibrium.

Figure 6.1, panel A, plots the evolution of the stock of H-1B visa holders from the I-Squared Act.[8] Year 0 is the estimated stock of H-1B visas under the current policy. The new policy is enacted in year 1. From steady state to steady state the stock about doubles from 510 thousand to 1.15 million. The transition does take time to play out, about ten years. This is due partly to the natural addition of new visas at the higher limit. The escalators also play an important role in slowing down the transition.

Figure 6.1, panel B, plots the evolution of the stock of all high-skilled labor. This includes new H-1B visas and the new mass of H-1B visa holders that transition to permanent status. It is normalized to 1 in year 0. When in enacted in year 1, the stock of high-skilled labor increases by a little less than .10 of a percent. Fifteen years out, the I-Squared Act leads to a 2 percent increase in the stock of high-skilled labor. Approximately 1 percentage point of the 2 percent is just from an expansion of the number of visas. The remaining 1 percentage point is from the increases in the flow of transitions to permanent status. Steady state to steady state, this policy leads to a 6 percent increase in the stock of high-skilled labor.

6.5.1   Relative Wages and Wage Elasticities

Measuring Changes in Relative Wages. In discussing the distributional effects of this policy, I focus on the measured elasticity of relative labor supply with respect to wages. I compute this measure by dividing the log change in relative labor supply by the log change in relative wages:

$$(23) \qquad \hat{\theta}_t = \frac{d\log(L_{st}) - d\log(L_{ut})}{d\log(w_{st}) - d\log(w_{ut})}.$$

I call this $\hat{\theta}_t$. Per Proposition 1, this is an interesting statistic because $\hat{\theta}_t$ and how it evolves reveals the extent to which firm dynamics and the complementarity between productivity and skill matter.

To understand this point, note that the estimator in equation (23) will recover the structural parameter θ, if there is no interaction between the skill

---

8. As an additional detail, I evaluate the policy with the proposed escalators described in the Immigration Innovation Act of 2015; that is, the number of H-1B visas increase only by 20,000 per year until the cap of 195,000 is reached.

A

B
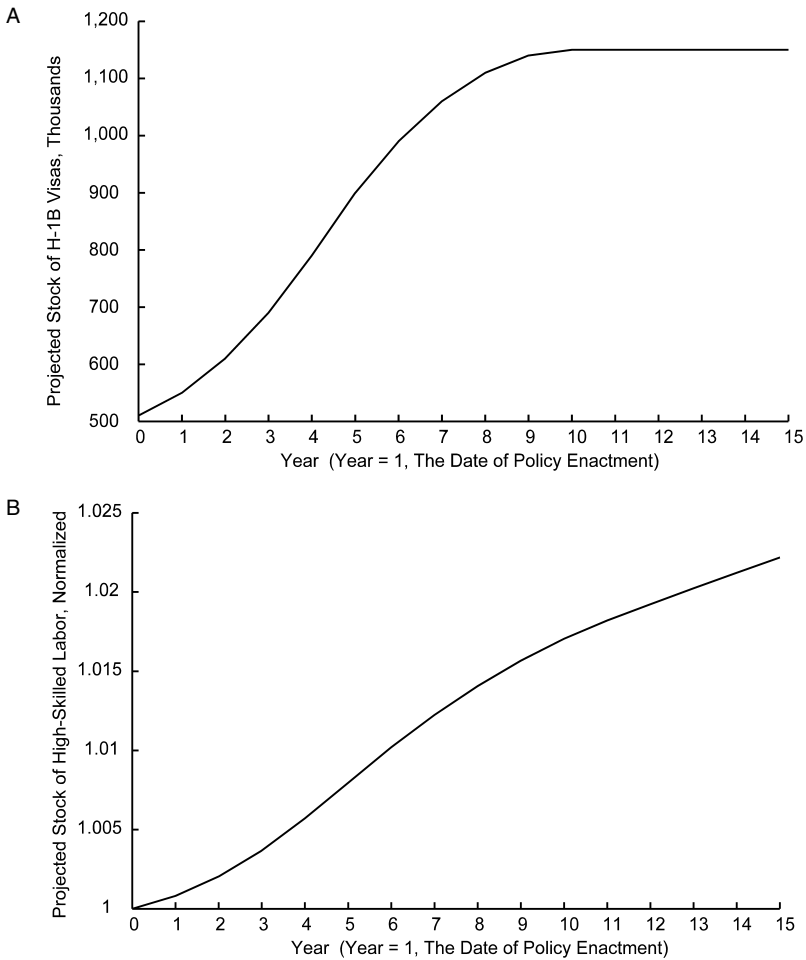
Fig. 6.1   Projected stocks of labor under the I-Squared Act

Notes: A, projected stock of H-1B visas; *B*, projected stock of high-skilled labor (normalized).

of the worker and the productivity of the firm. The calibrated model, how-
ever, finds a nontrivial amount of complementarity between high-skilled
workers and firm productivity (see the last row of table 6.1). Thus, Proposi-
tion 1 tells us that equation (23) will (a) deviate from the structural parameter
θ, and (b) vary over time. Thus, plotting $\hat{\theta}_t$ and how it evolves reveals the new
insights that the model can deliver about the change in relative wages.

Results: Wage Elasticities. Figure 6.2 plots $\hat{\theta}_t$. Year 1 is the date of policy
enactment; I plot this statistic going out only fifteen years. The dotted line
plots the elasticity in the long run—that is, the wage response as the economy
converges to the new stationary distribution.

Figure 6.2 shows that the wage elasticity is not constant and varies as the
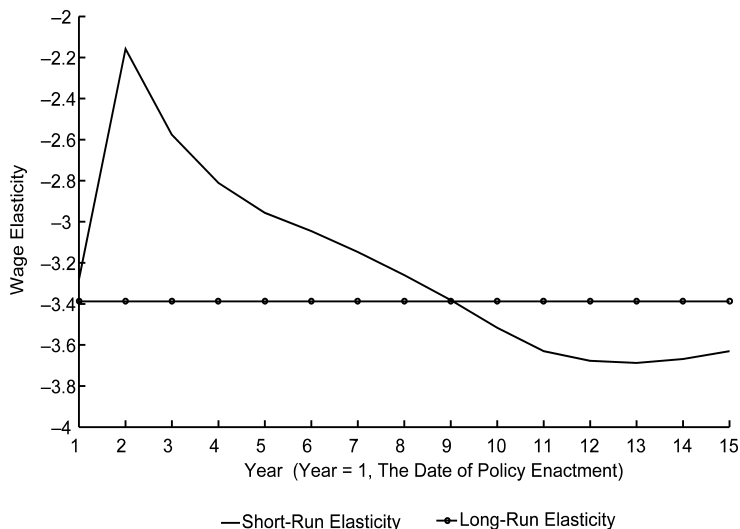
Fig. 6.2   I-Squared policy: Short- and long-run wage elasticities

change in policy plays out.[9] In the first year after the change in policy (year 2), the wage elasticity spikes at −2.2 and then gradually declines. That is, the skill premium shrinks more than the calibrated elasticity of substitution between skill types of −3 would imply. As the policy plays out, the wage elasticity undershoots and then converges to the dotted line of about −3.4.

There are several layers behind the explanation of the wage dynamics. Let me walk through the explanation in in steps.

First, the driving force is that new firms enter in response to the current and expected increases in high-skilled labor.[10] I plot the mass of entering firms in figure 6.3. I conjecture that the key reason is a market-size effect.[11] The size of the market expands, and, thus, entry takes place to bid down the returns of operating in the market and to equalize the free-entry condition in equation (13). This is analogous to variety expansion effects emphasized in monopolistic-competition models in Krugman (1980) or Melitz (2003).

Firm entry, however, is not sufficient to generate the dynamics in figure

9. The slight "bulge" between years 5 and 11 corresponds with when the escalators come off and the growth in stock of high-skilled labor accelerates slightly (see figure 6.1, panel B).

10. An interpretation of firm entry is that this is a form of "product innovation" in the language of Atkeson and Burstein (2010) and, thus, meshes well with the evidence in Kerr and Lincoln (2010).

11. I suspect other mechanisms are also at work. In particular, it raises the option value of entering. This policy makes high-productivity firms relatively more profitable, as the factor that they are using intensively has become more abundant. Firms stay in the market only if they have sufficiently high productivity. Thus, the downside (exit) is the same and the upside is more beneficial, and, hence, the option value of entering increased.
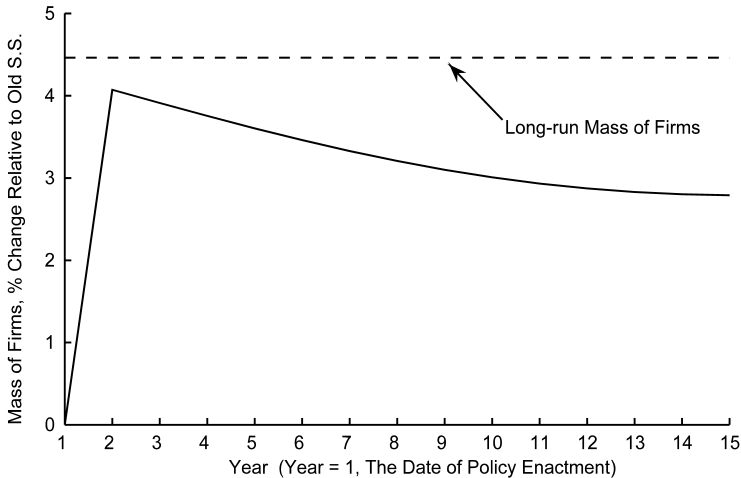
Fig. 6.3   **I-Squared policy: Mass of firms (relative to old SS)**

6.2. Proposition 1 says that there must be some form of skill bias across firms. Thus, the dynamics in figure 6.2 come from the interaction of firm entry and the skill bias across firms.

The intuition for how this interaction works is the following. First, new firms are likely to be low-productivity firms for two reasons: (a) entrants are not selected, as they come from an unconditional distribution, and (b) that unconditional distribution is also worse (the $\mu_e < 0$). Second, low-productivity firms use low-skilled labor more intensively. Thus, the expansion of low-productivity firms through entry bids up low-skilled wages more than would be expected. Thus, the skill premium decreases more than predicted by a constant-elasticity model.[12]

Changing the properties of the entry distribution and how the wage elasticity varies illustrates this point. For example, if $\mu_e = 0$, then new entrants will not be as unproductive relative to incumbents.[13] Thus, entry should not cause additional wage pressure for low-skilled workers to lead to a less responsive elasticity. This is exactly what figure 6.4 shows. The dashed line reports the wage elasticity when $\mu_e = 0$; the skill premium displays less dramatic dynamics.

12. The intuition here is closely related to the results of Burstein and Vogel (forthcoming) and their skill-biased productivity mechanism in response to trade liberalizations. The key distinction is the focus on the dynamic effects of a supply shock (immigration) rather than on a demand shock (opening to trade). Furthermore, my effects are driven by entry where as the model of Burstein and Vogel (forthcoming) has a fixed mass of firms.

13. This does not imply that entrants look like incumbents. Incumbents will be positively selected, as there is endogenous exit. Thus, even in this case, entrants will be less productive and demand relatively more low-skilled workers.
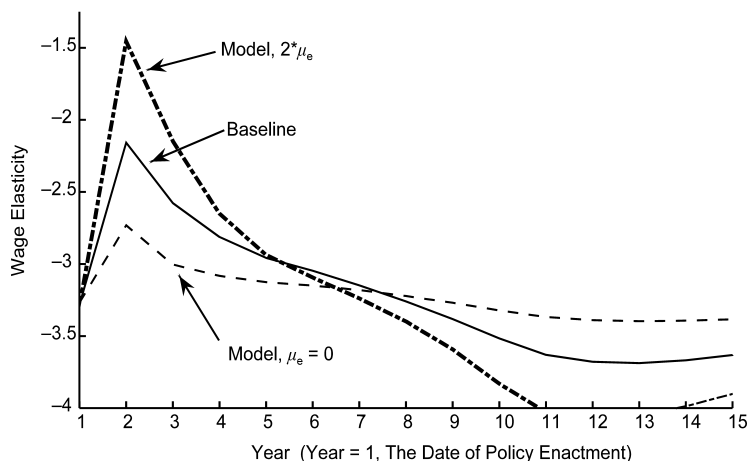
Fig. 6.4   Wage elasticities, baseline and alternative entry distributions

The corollary is that if entrants are even more (relative to the calibrated model) unproductive relative to incumbents, then the wage elasticity should vary more. Why? Entrants will be very unproductive, demand relatively more low-skilled workers, and place even more pressure on wages for low-skilled workers, leading to a more responsive elasticity. Again, this is exactly what figure 6.4 shows. The dash-dot line reports the wage elasticity when the shift in the entry distribution is twice its calibrated value $2*\mu_e$; the skill premium displays more dramatic dynamics.

To summarize: figure 6.2 shows that the skill premium contracts—and much more than a standard, constant-elasticity model would predict. The reason is that immigration makes the size of the market larger (today and in the future) and, thus, entry occurs. The calibrated models find that entrants are less productive and are low-skill intensive. Thus, entry bids up the relative price of low-skilled labor, and the skill premium decreases by more than a standard model would predict. The strength of this response depends on how different entrants are relative to incumbents.

Evidence Supporting the Mechanism. There are two aspects of the mechanism behind the results in figure 6.2: (a) firm entry responds to a change in labor supply, and (b) new firms are likely to be low productivity and low-skill intensive. There is evidence supporting both aspects of the mechanism.

First, research finds that changes in labor supply affect firm entry. In the context of changes in immigration, Olney (2013) presents compelling evidence in support of this piece of the mechanism; in US data, he finds a strong correlation between immigration and the new entry of establishments at the MSA level. In German data, Dustmann and Glitz (2015) show that firm entry and exit make important contributions to the absorption

of labor supply shocks. Karahan, Pugsley, and Şahin (2016) explore how demographic changes effect firm entry; using cross-state and industry data, they find that demographic changes have a large effect on the start-up rate of firms.

Second, there is evidence new firms are likely to be low productivity. To match the high exit rate of new firms, the model finds that new firms are less productive than the average incumbent is. This fact has been well documented (see, e.g., Baily, Hulten, and Campbell (1992); Bartelsman and Doms 2000).

What about how a firm's skill intensity varies with its productivity? Bernard and Jensen (1995) show that exporters (who are larger and more productive) pay higher wages relative to nonexporters. Thus, this suggests that high-productivity firms (exporters) demand more skilled workers and, hence, pay (on average) higher average wages. Schank, Schnabel, and Wagner (2007) discuss a similar finding in German data but establish that observable worker characteristics (e.g., education) account for the wage premium of exporters. This latter fact is very much in line with the calibration result that high-productivity firms employ relatively more high-skilled workers.

Burstein and Vogel (forthcoming) provide multiple pieces of evidence in support of the relationship between skill intensity and productivity. One compelling piece of evidence is that they find using the March CPS that the share of workers with a college degree is larger in larger firms in the United States. Using the correspondence between size and productivity in the model, this implies that high-productivity firms employ relatively more high-skilled workers.[14]

6.5.2   Wage Levels, Output, Consumption

The results in figure 6.2 show that the skill premium is shrinking—and shrinking more than the calibrated elasticity of substitution would imply. The temptation is to jump to the normative conclusion that high-skilled workers are worse off because of the I-Squared policy. The value added of a completely specified model is that I can evaluate the level effects on wages, output, and consumption. And, under certain conditions about the distribution of profits, a welfare evaluation.

Figure 6.5, panel A, plots the level of high- and low-skilled wages after the policy enactment. Again, year 1 is the date of the policy enactment; I plot statistics going out only fifteen years. In year 1, the level of high-skilled wages declines by −0.011 percent—essentially zero. In year 2, high-skilled workers' wages *increase* by 0.02 percent relative to prepolicy

---

14. Another piece of evidence builds on the observation in footnote 4. This model generates dispersion in TFPR and arises from differences in a firm's wage bill. Fox and Smeets (2011) use matched employer-employee panel data and find that (a) adjusting for labor quality reduces dispersion in TFPR, and (b) a firm's wage bill summarizes well the contribution of labor-quality differences to dispersion in TFPR.

A

B
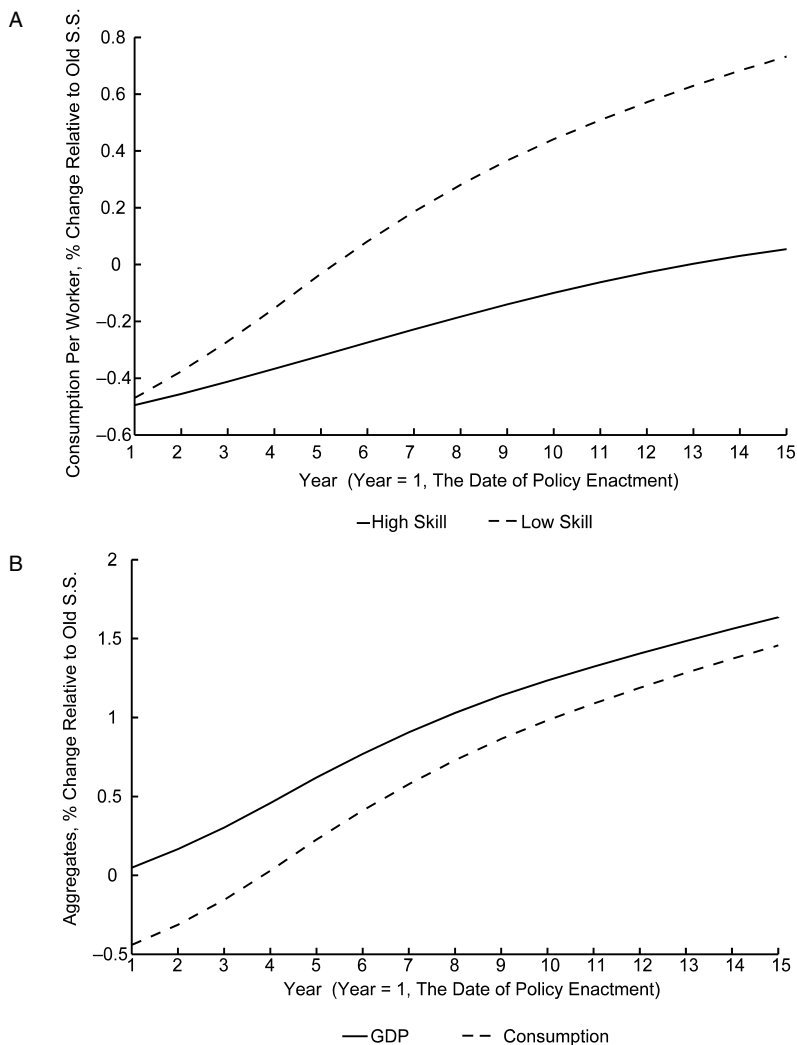
Fig. 6.5   I-Squared policy: Wage levels, output, and consumption
Note: A, I-Squared; *B*, I-Squared policy: output and consumption, percent change relative to old steady state.

levels—small, but not negative. Thus, one year after the policy, high-skilled workers are better off (in terms of labor earnings) than before the enactment of the policy.

Scale and productivity effects are the reasons high-skilled workers earn more. First, a larger labor force leads to more firms (product variety) and higher real wages for all. This expansion in the mass of firms is clearly seen in figure 6.3 with a 4 percent increase in the mass of firms in year 1. Second,

there is a productivity effect. More high-skilled labor at a relatively lower price allows firms to substitute into more productive labor, and this is happening for the most productive firms. The analogue to the trade literature is that opening to trade reallocates resources toward the most productive firms, as in Melitz (2003).

The entry of firms is typically thought of as a long-run effect (see, e.g., di Giovanni, Levchenko, and Ortega 2015). However, this benefit is felt in year 2—not in the "long run." The reason is the dynamic, forward-looking nature of the firm. In year 1, the mass of firms increases by 4 percent even though the stock of high-skilled labor expands by only 0.20 percent at that point. Thus, firms respond forcefully to the current and foreseen increases in labor supply.

These effects show up in GDP. Figure 6.5, panel B, plots the path of aggregate output (solid line). In year 1, there is a 0.05 percent increase in GDP. After the complete transition, the I-Squared Act delivers a 4.5 percent increase in GDP, given a 2 percent increase in the total stock of labor.

These benefits are not without costs—and these costs show up as a loss in consumption. The dashed line in figure 6.5, panel B, plots aggregate consumption. In year 1, there is nearly a −0.50 percent loss in consumption. It is not until year 4 that the level of consumption is at its prepolicy level.

The issue is that all the "good" outcomes—higher wages and higher output—come from firm entry. However, firm entry must be paid for. There is investment today in the creation of firms to prepare for a larger labor force in the future—and this investment comes at the cost of consumption. This result makes clear that near-zero wage impacts (e.g., figure 6.5, panel A) do not imply near-zero effects on economic outcomes that are closer a measure of welfare. Thus, the normative implications of immigration policy are more nuanced than just the change in wages.

I made some important modeling choices that might change the consumption result. The first one regards the lack of curvature over period utility in consumers' preferences in equation (1). If consumers have a finite intertemporal elasticity of substitution, then the interest rate firms' profits are discounted would be endogenous and depend on the path of consumption. A growing path of consumption would lead to an increase in the interest rate; firms would discount profits more heavily, and then firm entry would not be as strong. A related issue is the closed economy assumption—international borrowing would allow consumers to smooth consumption while these investments are made.

The second important modeling choice was the denomination of the entry and fixed costs in units of output versus the alternative being the denomination in units of labor. The upside of this choice was that (a) I avoided having to make choices regarding the skill intensity of these activities and then feeding into the distributional effects, and (b) it provides a long-run motive for immigration through the introduction of a scale effect. The downside is

that (a) it introduces a scale effect that may be empirically implausible, and (b) the increases in firm entry lead to losses in consumption as more investment in new firms takes place.

6.5.3   I-Squared Policy: Welfare

The I-Squared policy leads to essentially no negative wage effects; yet, aggregate consumption falls on impact and stays depressed for up to four years. This implies that the welfare effects depend critically on the distribution of firms' profits across workers. Below, I illustrate this issue by taking a stand on the distribution of profits and then discuss alternatives.

To compute welfare, I assume that consumers of a skill type receive a wage-bill weighted share of profits net of entry costs. This allocation rule implies that consumption of skilled and unskilled workers is

$$c_s = w_s + \Pi\left(\frac{w_s}{w_u L_u + w_s L_s}\right),$$

$$c_u = w_u + \Pi\left(\frac{w_u}{w_u L_u + w_s L_s}\right),$$

where Π is aggregate profits; and $w_s/(w_u L_u + w_s L_s)$ is the wage-bill share that a skilled worker receives.

Figure 6.6 plots consumption per worker. Consumption for both workers falls on impact and stays depressed, relative to the old steady state, for four and five years for low- and high-skilled workers. Figure 6.5, panel A, shows that labor income is little changed. Thus, consumption falls because the profit that both worker groups receive falls as new entry takes place. Because the workers are also the owners of the firms, they bear the cost of investment today for the creation of firms tomorrow.

As a formal welfare metric, I compute the present discounted gain in consumption for both worker types over the entire transition path into the infinite future. The I-Squared policy amounts to 1.29 and 0.22 percent increases in present discounted consumption for low- and high-skilled workers. This is a substantial—especially for such a seemingly small expansion in the H-1B visa program. With that said, it is also about 50 percent less than a simple, long-run evaluation that compares across steady states would suggest (2.46 percent and 0.65). Thus, the adjustment to this policy is an important consideration.

I do not believe the conjectured profit-sharing rule is accurate or realistic. But this thought experiment makes an important conceptual point: the owners of the firm, not the workers, bear the burden of the adjustment to the I-Squared policy. Mark Zuckerberg and Bill Gates—who are large shareholders of the firms that they operate (or operated) and who advocate policies such as the I-Squared Act—bear the short-run burden.
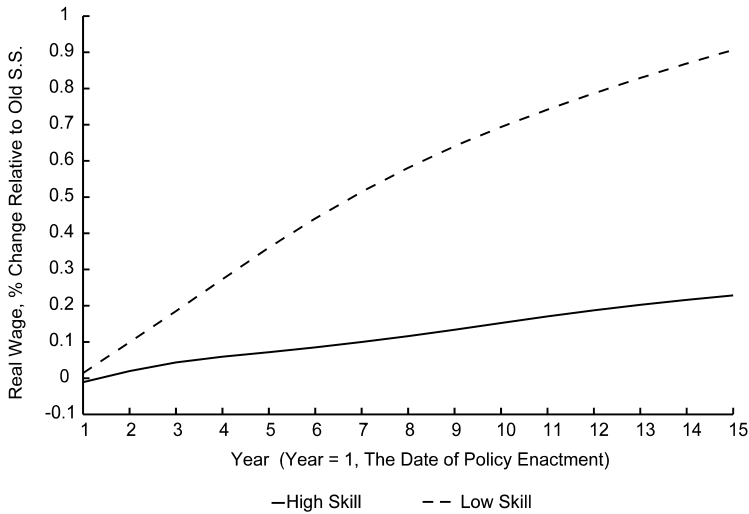
Fig. 6.6   I-Squared policy: Consumption per worker

6.6   Nationalistic Policy

Since I started writing the chapter, the policy environment in the United States has changed. Currently on the table are discussions about policies that are "nationalistic," in the sense that they restrict the movement of labor and goods with the goal of protecting national interests. This section evaluates the economic benefits/costs of one such policy: a complete elimination of the H-1B visa program.

To evaluate this policy, I build on the discussion in section 6.4.3 and change several things. First, no new H-1B visa holders are allowed in the United States. Thus, there is no longer a flow of 85,000 high-skilled workers into the economy each year. However, I do allow existing H-1B visas holders to remain for the maximum duration of six years. Finally, existing H-1B visa holders are prevented from changing their visa status and transitioning into permanent status—that is, H-1B transitions in equation (20) are set to zero. This is probably an extreme assumption, as it implies that nonemployment transitions into permanent status (e.g., marriage) are not possible.

Figure 6.7 plots the projected evolution of the stock of all high-skilled labor. It is normalized to 1 in year 0. The nationalistic policy is enacted in year 1, and the stock of high-skilled labor decreases by a little more than .10 of a percent—essentially, the 85,000 visa holders who exit and are not replaced. Over the next six years, the stock of high-skilled labor falls as H-1B visa holders are not replaced. The stock of high-skilled labor continues to decline as the inflow of domestic college graduates is not sufficient to
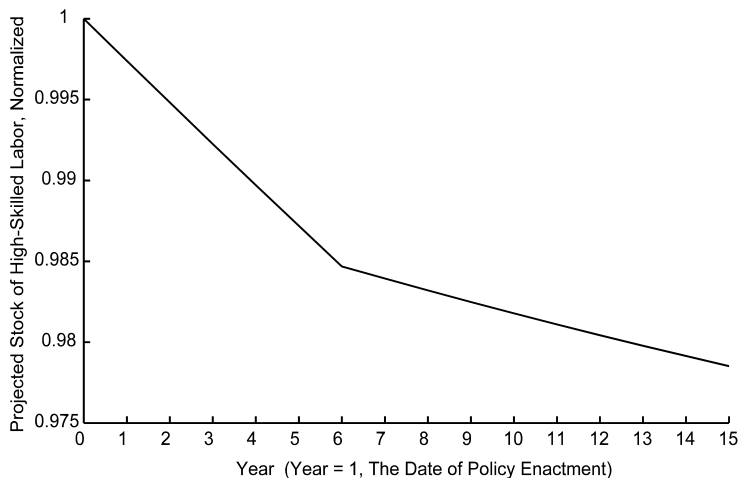
Fig. 6.7   Nationalistic policy: Projected stock of high-skilled labor (normalized)

replace the outflow of high-skilled workers. This latter statement is about the δ parameter, which was calibrated such that labor supply was stationary under current policy. Fifteen years out, the nationalistic policy leads to a 2 percent decrease in the the stock of high-skilled labor.

As in the previous analysis, I treat the closing of high-skilled immigration as unanticipated from the perspective of firms. After the policy is announced, firms understand the entire projected path of the workforce. I then compute the transition path of the economy to its new stationary equilibrium.

The next two subsections discuss the impact of this policy on the structure of wages and then on consumption, output, and welfare.

6.6.1   Nationalistic Policy: Wages

Figure 6.8, panel A, plots the level of high- and low-skilled wages after the policy enactment. Again, year 1 is the date of policy enactment; I plot this statistic going out only fifteen years. In year 1, the level of high-skilled wages increase by 0.04 percent—essentially zero. In year 2, high-skilled wages are unchanged, and by year 3 they decrease −0.02 percent relative to prepolicy levels. In the near term, reducing the supply of skilled workers does essentially nothing to increase their wages relative to before the enactment of the policy.

As discussed in the previous section, it is both a scale and productivity effect that leads to high-skilled workers now earning less. In this case, a smaller labor force leads to fewer firms and to lower real wages for all. Figure 6.8, panel B, shows how the mass of firms drops dramatically, with a 6 percent decrease in the mass of firms in year 2. Again, to emphasize this point, this is contrary to the conventional wisdom that the entry of firms is

A

B
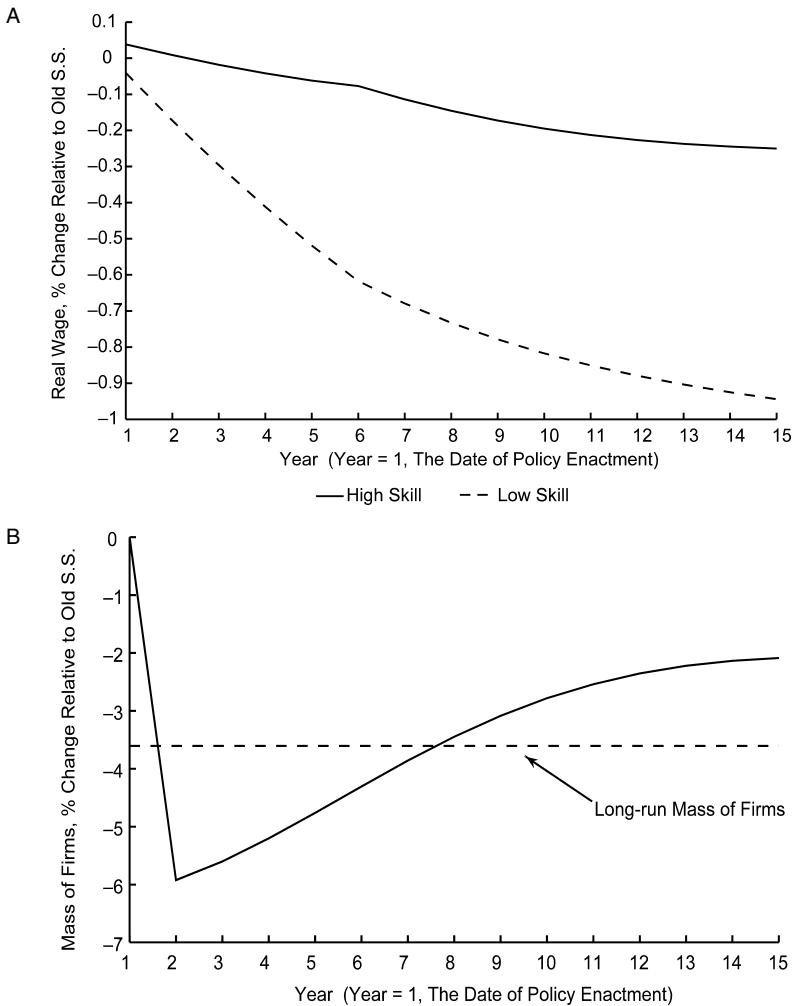
Fig. 6.8   Nationalistic policy: Wages and the mass of firms

Note: A, wage levels; *B*, mass of firms.

a long-run effect. Because of the dynamic, forward-looking nature of the firm, the long-run costs of a more restrictive immigration policy are quickly felt. Furthermore, the productivity effect comes from firms substituting into low-skilled labor as high-skilled labor is now relatively more expensive (discussed below).

   Restricting high-skilled immigration has unintended, negative consequences on the wages of low-skilled workers. The dashed line in figure 6.8, panel A, plots the level of low-skilled wages after the policy enactment; these drop relatively sharply in the near term and then more gradually over

the transition. The key reason for the drop in low-skilled wages (relative to high-skilled workers) is that the firms who are not entering anymore and the firms that are exiting are low-productivity and, hence, low-skill-intensive firms. In other words, the shrinkage of the market differentially lowers the demand for low-skilled labor because it is the low-skill-intensive firms that leave.

The negative impact on low-skilled wages further illustrates the mechanisms (but in the opposite direction) seen in the I-Squared policy. In this case, firms foresee a smaller market, which results in less entry and more exit, and entrants and exiting firms are less productive *and* low-skill intensive. Thus, the skill premium increases and does so at the expense of those workers at the bottom—not at the top—as the demand for their labor services decreases by relatively more.

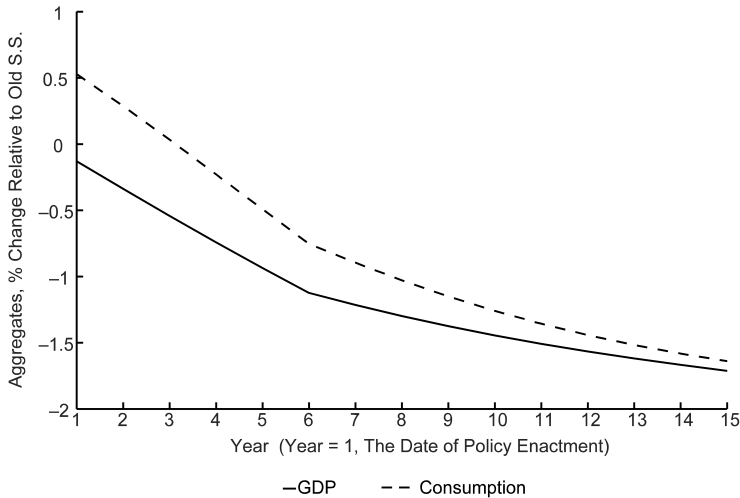### 6.6.2   Nationalistic Policy: Output, Consumption, Welfare

The reduction in the labor force and the mass of firms shows up as a drop in GDP. Figure 6.9 plots GDP (solid line) relative to its old steady-state value. In year 1, there is a 0.13 percent decrease in GDP. This decline continues as fewer firms enter, existing firms exit, and the labor force contracts. After the complete transition, the nationalistic policy delivers a 3.7 percent decrease in GDP.

These losses in wages and output do come with the benefit of a short-run gain in consumption. Figure 6.9 plots aggregate consumption relative to its old steady-state value. On impact, there is a 0.5 percent increase in consumption, but over time this effect dissipates, and consumption eventually declines in a similar manner as GDP.

In the nationalistic policy, lower wages and lower output come from firm entry, or lack thereof. What this means is that there is no need to invest in the creation of new firms because the economy has too many firms, given the reduction in the labor force (today and in the future). And, the reduction in investment comes at the benefit of higher consumption in the short run.

As in evaluating the I-Squared policy, the normative implications of the nationalistic immigration policy are more nuanced than just the change wages and depend on how profits are redistributed throughout the economy. Using the conjectured profit-sharing rule discussed in section 6.5.3, I compute the present discounted gain in consumption. The nationalistic policy amounts to a 0.20 decrease in present discounted consumption for high-skilled workers. For low-skilled workers—whom this policy is presumably intended not to affect—the welfare decrease is nearly six times as large, at a 1.13 percent decrease in present discounted consumption.

Interestingly, the consequences of the nationalistic policy are borne by the workers, not by the owners of the firms. The firms' owners benefit from this policy as they enjoy higher profits in the short run as a smaller market size reduces the necessity to invest in the creation and maintenance of firms.

Fig. 6.9   Nationalistic policy: GDP and consumption

6.7   Conclusion

This chapter places the firm and its dynamics at the center of economic issues regarding immigration—in particular, the distributional effects, aggregate outcomes, and welfare. This chapter provides a useful place to start, but questions remain.[15] To conclude, let me outline three key lessons and, within each lesson, some open empirical questions.

Firm Dynamics and the Distributional Impacts of Immigration. The dynamics of the firm at the micro level generate nontrivial, short-run dynamics in relative wages that differ from their long-run dynamics. An expansion of skilled labor leads the wage premium of high-skilled to low-skilled workers to shrink more than a standard, static CES model would predict. Similarly, a contraction of skilled labor leads to an increase in the skill premium by more than the standard model would predict. Crucial to delivering this result is entering firms' skill bias relative to incumbent firms. Understanding this last point and what the data say about it is important. This result also raises, perhaps, questions about the structural interpretation of empirical evidence on the wage response to immigration.

The Aggregate Impacts of Immigration. In the chapter's evaluation of the I-Squared Act, wages essentially never declined and there was a 1.5 percent increase in GDP after fifteen years. The nationalistic policy quickly led to

15. Kerr, Kerr, and Lincoln (2014) provide a nice summary of the many open questions regarding firms and immigration. Xu (2016) moves in this same direction by focusing on the impact of high-skilled immigration in a growth model with firm dynamics in which firms engage in innovative and growth-enhancing activity.

declines in wages for both skill groups and a 1.5 percent decrease in GDP after fifteen years.

There are two elements to these outcomes. First, these effects came from firms entering quickly in response to the policy. In other words, labor demand shifted and did so quickly. An open question regards the elasticity of firm entry to labor supply. Olney (2013), Dustmann and Glitz (2015), and Karahan, Pugsley, and Şahin (2016) provide suggestive evidence, but this is an important detail that deserves more scrutiny. Second, the aggregate gains/losses partially depend on a scale effect in the model. Scale effects are hard to identify in the data, but (as in fields such as trade and growth) they are critical to evaluating the gains from immigration.

Who Bears the Burden of Adjustment? The expansion of immigration leads to gains and its contraction leads to losses. But how these gains are distributed depends less on its effect on labor earnings and more on distribution of profits. In the I-Squared policy, it was the owners of the firms that bore the burden. Under the nationalistic policy, the owners of the firm reaped short-run gains, while workers (low-skilled workers in particular) lost. Essentially, these observations place the wealth distribution at the center of the analysis. While perhaps obvious to me now, this came as a surprise, and it opens the door to many interesting questions about the winners and losers from changes in immigration and the distribution of wealth.

Appendix A

Data

This section describes the data used to calibrate the model.

1. Autocorrelation of establishment size. These moments were computed using the Synthetic Longitudinal Database (US Census Bureau 2011) by regressing the logarithm of establishment size across consecutive years for those establishments that are present in both years. Only establishments with more than one employee were used. This procedure yields an autocorrelation coefficient of 0.90, which I found to be stable across years and unaffected by incorporating industry fixed effects (at the SIC3 level).

2. Entry and survival rates. These moments were computed using the Synthetic Longitudinal Database (US Census Bureau 2011). The entry rate is computed as the new establishments relative to the total number of establishments. This number is computed to be about 10 percent in the later time periods of the data set. The survival rate is computed as establishments starting in a given-year period that remain open (over some time horizon) relative to all establishments starting in that year. This number is about 50 percent at a five-year horizon.

3. Fraction of skilled to unskilled workforce and earnings. These moments

were computed using the Current Population Survey. Unskilled workers were computed as the sum of the labor force with less than a bachelor's degree (series id: LNS11027659, LNS11027660, LNS11027689). Skilled workers were computed as those with a bachelor's degree or more (series id: LNS11027662). To abstract from long-run trends in the skill composition of the labor force, I focused on the average ratio of skilled to unskilled workers for the time period from 2010 to 2015; this led to an estimate of 0.57.

The relative earnings of skilled to unskilled workers were computed using the Current Population Survey. Using the same definition as above, I compared median usual weekly earnings for skilled and unskilled workers (series id: LEU0252916700, LEU0252917300, LEU0254929400, LEU0252918500). This provides an estimate that skilled workers earn 1.89 times as much as unskilled workers.

4. Firm size and size-wage premium. I used the Statistics of US Businesses from the US Census Bureau. The Statistics of US Businesses reports data that include firms binned by size with data on the number of firms, the number of establishments, employment, and the annual payroll for most US businesses. To compute a measure of the dispersion in firm size, I find that half of all employment is in firms with more than 500 employees; I abuse terminology here, but I will call this the median firm. The average firm size is about twenty employees. Thus, I target a ratio of the median to mean size of twenty-five.

The size-wage premium is determined as follows. I compute the payroll divided by employment for those firms with more than 500 employees—I call this the average wage above the median. Then, I compare this to the average wage or workers in firms below 500 employees. For the period from 2010 to 2013, this value is 1.30. That is, the average wage in firms with more than 500 employees is 30 percent larger than in firms with fewer than 500 employees.

Appendix B

Computing the Transition Path

This section describes how to compute the deterministic transition associated with an unexpected change in a primitive. The basic idea is to guess a sequence of endogenous values and (a) work backward solving for the policy function of the firm, then (b) solve the distribution of firms going forward using the policy function, and (c) check that the guessed endogenous values are consistent with market clearing implied by (b). Finally, update the conjectured about the endogenous variables in a smart way until market-clearing conditions are satisfied.

1. First, hand the computer several things: the value function associated

with the new stationary equilibrium (call it $v_T(z,\mathbf{p_T})$), and an initial distribution of firms (call it $\mu_0(\mathbf{p_0})$) that describes the mass of continuing and new firms at the end of the period, just prior to the change in the labor force.

2. Guess a sequence of $\{\mathbf{w}_t, Y_t, M_t\}_{t=1}^{T}$ where T is the end point that corresponds to the new stationary equilibrium; $\mathbf{w}$ is the vector of wages per efficiency units for each skill type; Y_t is aggregate output; M_t is the mass of entering firms. To economize on notation, denote this sequence as $\{\mathbf{p}_t\}_{t=1}^{T}$.

3. Given $\{\mathbf{p}_t\}_{t=1}^{T}$ and $v_T(z_j, \mathbf{p_T})$, work backward to compute the value and policy functions where the explicit dependence on $\mathbf{p}$ is made. So, the value of the firm must respect

(B.1) $v_{T-1}(z_i, \mathbf{p_{T-1}}) = \max\left[\pi(z_i, \mathbf{p_{T-1}}) - \kappa + \beta \sum_{j=1}^{m} \mathscr{P}(i,j) v_T(z_j, \mathbf{p_T}), \pi(z_i, \mathbf{p_{T-1}}) \right].$

Then, as we walk this backward, this generates a sequence of policy functions—call it $g(\mathbf{p_t})$ for each date t.

4. Given the policy functions, take the initial distribution of firms $\mu_0(\mathbf{p_0})$ and solve forward to compute how the mass of firms evolves given the policy functions that solve equation (B.1)

(B.2) $$M_t \tilde{\mathscr{P}}_e + \mu_t * \left(P \circ g(\mathbf{p_t}) \right) = \mu_{t+1},$$

which yields a sequence of measures over firms for every date $\{\mu_t(\mathbf{p_t})\}_{t=1}^{T}$.

5. Given the sequence of measures over firms, check whether markets clear and the free-entry condition is satisfied at each date t. A second condition is that, at date T, the economy should have converged to the new stationary equilibrium. Thus, check whether $v_{T-1}(z_i, \mathbf{p_{T-1}})$ is close to the value function $v_T(z_j, \mathbf{p_T})$ associated with the new stationary equilibrium.

6. If these conditions are not satisfied, update the guessed sequence of prices $\{\mathbf{w}_t, Y_t, M_t\}_{t=1}^{T}$ and work until they are met.

References

Alessandria, George, Horag Choi, and Kim Ruhl. 2014. "Trade Adjustment Dynamics and the Welfare Gains from Trade." NBER Working Paper no. 20663, Cambridge, MA.

Atkeson, Andrew, and Ariel Burstein. 2010. "Innovation, Firm Dynamics, and International Trade." *Journal of Political Economy* 118 (3): 433–84.

Baily, Martin Neil, Charles Hulten, and David Campbell. 1992. "Productivity Dynamics in Manufacturing Plants." *Brookings Papers on Economic Activity: Microeconomics* 1992:187–267.

Bartelsman, Eric J., and Mark Doms. 2000. "Understanding Productivity: Lessons from Longitudinal Microdata." *Journal of Economic Literature* 38 (3): 569–94.

Bernard, Andrew, Jonathan Eaton, J. Bradford Jensen, and Samuel Kortum. 2003. "Plants and Productivity in International Trade." *American Economic Review* 93 (4): 1268–90.

Bernard, Andrew B., and J. Bradford Jensen. 1995. "Exporters, Jobs, and Wages in US Manufacturing: 1976–1987." *Brookings Papers on Economic Activity* 1995:67–119.

———. 1999. "Exceptional Exporter Performance: Cause, Effect, or Both?" *Journal of International Economics* 47 (1): 1–25.

Borjas, George J. 2003. "The Labor Demand Curve Is Downward Sloping: Reexamining the Impact of Immigration on the Labor Market." *Quarterly Journal of Economics* 118 (4): 1335–74.

———. 2014. *Immigration Economics.* Cambridge, MA: Harvard University Press.

Broda, Christian, and David E Weinstein. 2006. "Globalization and the Gains from Variety." *Quarterly Journal of Economics* 121 (2): 541–85.

Burstein, Ariel, and Jonathan Vogel. Forthcoming. "International Trade, Technology, and the Skill Premium." *Journal of Political Economy*.

Caliendo, Lorenzo, and Fernando Parro. 2015. "Estimates of the Trade and Welfare Effects of NAFTA." *Review of Economic Studies* 82 (1): 1–44.

Card, David. 2009. "Immigration and Inequality." *American Economic Review* 99 (2): 1–21.

Chaney, Thomas. 2008. "Distorted Gravity: The Intensive and Extensive Margins of International Trade." *American Economic Review* 98 (4): 1707–21.

Decker, Ryan, John Haltiwanger, Ron Jarmin, and Javier Miranda. 2014. "The Role of Entrepreneurship in US Job Creation and Economic Dynamism." *Journal of Economic Perspectives* 28 (3): 3–24.

di Giovanni, Julian, Andrei A. Levchenko, and Francesc Ortega. 2015. "A Global View of Cross-Border Migration." *Journal of the European Economic Association* 13 (1): 168–202.

Dustmann, Christian, and Albrecht Glitz. 2015. "How Do Industries and Firms Respond to Changes in Local Labor Supply?" *Journal of Labor Economics* 33 (3, part 1): 711–50.

Foster, Lucia, John Haltiwanger, and Chad Syverson. 2008. "Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?" *American Economic Review* 98 (1): 394–425.

Fox, Jeremy T., and Valerie Smeets. 2011. "Does Input Quality Drive Measured Differences in Firm Productivity?" *International Economic Review* 52 (4): 961–89.

Grogger, Jeffrey, and Gordon H. Hanson. 2011. "Income Maximization and the Selection and Sorting of International Migrants." *Journal of Development Economics* 95 (1): 42–57.

Hathaway, Ian, and Robert E. Litan. 2014. "Declining Business Dynamism in the United States: A Look at States and Metros." *Brookings.* https://www.brookings.edu/research/declining-business-dynamism-in-the-united-states-a-look-at-states-and-metros/.

Hopenhayn, Hugo A. 1992. "Entry, Exit, and Firm Dynamics in Long Run Equilibrium." *Econometrica: Journal of the Econometric Society* 60 (5): 1127–50.

Hopenhayn, Hugo, and Richard Rogerson. 1993. "Job Turnover and Policy Evaluation: A General Equilibrium Analysis." *Journal of Political Economy* 101 (5): 915–38.

Hsieh, Chang-Tai, and Peter J Klenow. 2009. "Misallocation and Manufacturing TFP in China and India." *Quarterly Journal of Economics* 124 (4): 1403–48.

Karahan, Fatih, Benjamin Pugsley, and Ayşegül Şahin. 2016. "Demographic Origins of the Startup Deficit." Working Paper, Federal Reserve Bank of New York. https://economics.nd.edu/assets/217341/pugsley_demographics_startups_1_.pdf.

Kerr, Sari Pekkala, William R. Kerr, and William F. Lincoln. 2014. "Firms and the Economics of Skilled Immigration." NBER Working Paper no. 20069, Cambridge, MA.

Kerr, William R., and William F. Lincoln. 2010. "The Supply Side of Innovation: H-1B Visa Reforms and US Ethnic Invention." *Journal of Labor Economics* 28 (3): 473–508.

Krugman, Paul. 1980. "Scale Economies, Product Differentiation, and the Pattern of Trade." *American Economic Review* 70 (5): 950–59.

Krusell, Per, Lee E. Ohanian, José-Víctor Ríos-Rull, and Giovanni L. Violante. 2000. "Capital-Skill Complementarity and Inequality: A Macroeconomic Analysis." *Econometrica* 68 (5): 1029–53.

Lee, Hyun. 2016. "Quantitative Impact of Reducing Barriers to Skilled Labor Immigration: The Case of the U.S. H-1B Visa." Working Paper no. 2016-35, Department of Economics, University of Connecticut.

Lewis, Ethan. 2011. "Immigration, Skill Mix, and Capital Skill Complementarity." *Quarterly Journal of Economics* 126 (2): 1029–69.

———. 2013. "Immigration and Production Technology." *Annual Review of Economics* 5 (1): 165–91.

Lowell, B. Lindsay. 2000. "H-1B Temporary Workers: Estimating the Population." CCIS Working Paper, Center for Comparative Immigration Studies.

Melitz, Marc. 2003. "The Impact of Trade on Aggregate Industry Productivity and Intra-Industry Reallocations." *Econometrica* 71 (6): 1695–725.

Olney, William W. 2013. "Immigration and Firm Expansion." *Journal of Regional Science* 53 (1): 142–57.

Ottaviano, Gianmarco I. P., and Giovanni Peri. 2012. "Rethinking the Effect of Immigration on Wages." *Journal of the European Economic Association* 10 (1): 152–97.

Parro, Fernando. 2013. "Capital-Skill Complementarity and the Skill Premium in a Quantitative Model of Trade." *American Economic Journal: Macroeconomics* 5 (2): 72–117.

Ruhl, K. J. 2008. "The International Elasticity Puzzle." Unpublished Manuscript, Penn State University.

Schank, Thorsten, Claus Schnabel, and Joachim Wagner. 2007. "Do Exporters Really Pay Higher Wages? First Evidence from German Linked Employer-Employee Data." *Journal of International Economics* 72 (1): 52–74.

Simonovska, Ina, and Michael E. Waugh. 2014a. "The Elasticity of Trade: Estimates and Evidence." *Journal of International Economics* 92 (1): 34–50.

———. 2014b. "Trade Models, Trade Elasticities, and the Gains from Trade." NBER Working Paper no. 20495, Cambridge, MA.

Spreen, Thomas Luke. 2013. "Recent College Graduates in the US Labor Force: Data from the Current Population Survey." *Monthly Labor Review* 136 (3): 3–12. https://www.bls.gov/opub/mlr/2013/02/art1full.pdf.

Tauchen, G. 1986. "Finite State Markov-Chain Approximations to Univariate and Vector Autoregressions." *Economics Letters* 20 (2): 177–81.

US Census Bureau. 2011. Synthetic LBD Beta version 2.0. [computer file]. Washington, DC and Ithaca, NY: US Census Bureau and Cornell University, Synthetic Data Server [distributor].

Xu, Rui. 2016. "High-skilled Migration and Global Innovation." SIEPR Discussion Paper no. 16-016, Stanford Institute for Economic Policy Research, Stanford University. https://siepr.stanford.edu/sites/default/files/publications/16-016.pdf.