The dynamics of ethnic segregation in the German labour market*

Sébastien Willis†

December 2022

Abstract

Segregation in residential markets, labour markets, or school systems is often associated with preference spillovers and the presence of tipping points in group composition. I use administrative data from Germany to study whether such tipping dynamics are also observed in the composition by nativity of firms. I find only limited evidence of tipping points; the evidence is strongest for firms in relatively low-skill industries during years of high immigrant inflows. This suggests that workplace segregation is unlikely to be explained in general by the preferences of workers over the composition of their workplaces. However, evidence on the dynamics of aggregate segregation suggests that preference spillovers may become relevant to understanding changes in segregation in certain industries when immigrant inflows are sufficiently large.

Keywords: Segregation, firms, tipping points, immigration
JEL codes: J15, J61

*An earlier version of this paper was the first chapter of my dissertation at Universitat Pompeu Fabra. I would like to thank my advisors, Albrecht Glitz and Ruben Enikolopov, as well as Caterina Alacevich, Lídia Farré, Flavio Hafner, Linna Martén, Joan Monras, Dávid Nagy, Ana Rodríguez González, Matti Sarvimäki, anonymous referees, and seminar participants at the UPF applied seminar, CREI international lunch, and Uppsala University for helpful feedback. Thanks also to Adam McCloskey for helpful discussions about implementing the inference methods in Andrews et al. (2021) and to Heiner Frank and the staff at the IAB for help accessing the data.

†Department of Economics, Uppsala University, sebastien.willis@nek.uu.se. This paper uses data from the Establishment History Panel 1975–2019, v2 (doi: 10.5164/IAB.BHP7519.de.en.v2), available upon application to the Institute for Employment Research. There are no relevant conflicts of interest or sources of funding to declare.

1 Introduction

Immigrants make up an increasingly large share of the workforce in developed countries and Germany has been no exception, with the foreign-born representing 16.1 per cent of the German population in 2019 (OECD, 2020). However, once they enter the labour market, immigrants and natives tend not to work for the same firms. In 2008, when immigrants already made up 12.8 per cent of the population (OECD, 2020), 40 per cent of immigrants in Germany would have needed to change firms to achieve a degree of workplace segregation consistent with a random assignment of workers to firms (Glitz, 2014). Workplace segregation has also been document in other high-immigration countries including the US (Andersson et al., 2014; Hellerstein and Neumark, 2008) and Sweden (Åslund and Skans, 2010).[1]

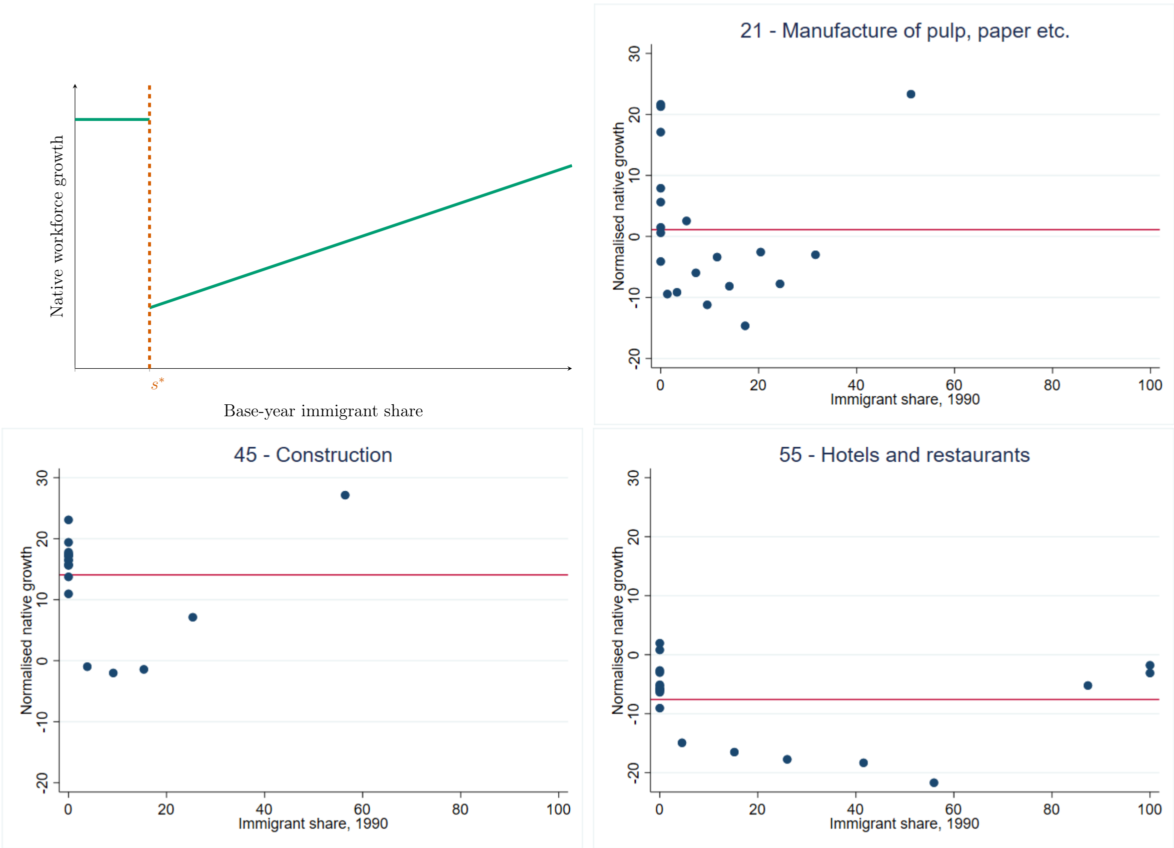
While there is ample cross-sectional evidence of segregation in the workplace, our understanding of the causes of workplace segregation is limited by a relative lack of empirical evidence on the dynamics of segregation across firms over time. This is in contrast to other settings where segregation has been documented, and in particular residential neighbourhoods, where the dynamics of segregation have been extensively studied. There, a long-standing literature has considered the role of endogenous feedback from past neighbourhood composition to future changes in neighbourhood composition. The possibility that such feedback loops might lead to tipping points in the composition of neighbourhoods has been considered both theoretically (Schelling, 1971, 1978; Becker and Murphy, 2000; Banzhaf and Walsh, 2013) and empirically (Aldén et al., 2015; Caetano and Maheshri, 2021; Card et al., 2008, 2011).

In this paper I study whether such tipping dynamics also exist in the composition of workplaces. Some examples of tipping dynamics are given in Figure 1. In the top-left panel, I provide a stylised example of a tipping dynamic. For low values of the immigrant share, native workforce growth is unrelated to the base-year immigrant share. However, should the immigrant share cross some threshold, labelled s^* and referred to as a tipping point, native workforce growth falls sharply, potentially becoming negative, as the firm's hiring shifts towards immigrants. In the remaining panels, I report a measure of native workforce growth for three two-digit industries—Manufacture of pulp and paper, Construction, and Hotels and Restaurants—during the period 1990–1995, a period of high immigration to Germany. For this selected sample of industries, we see evidence of tipping-like dynamics. Firms with low initial immigrant shares experience above-average native

[1] Workplace segregation unexplained by observed characteristics suggests factors of production are misallocated, which could have large negative consequences for aggregate productivity and output (Hsieh et al., 2019). At the individual level, segregation across workplaces or industries could help explain the widely-studied persistence of employment and wage gaps between immigrants and natives (e.g. Lubotsky, 2007; Rho and Sanders, 2021) and the fact that immigrants tend to work at lower-paying firms (Aydemir and Skuterud, 2008; Barth et al., 2012).

workforce growth, which then decreases precipitously for somewhat larger values of the initial immigrant share, before recovering for even higher values of the initial immigrant share.

Figure 1: Examples of tipping dynamics



Notes: Normalised native workforce growth (defined in the text) as a function of the immigrant share in the base year, either in an abstract example, or in actual industries in 1990–1995. When plotting actual data, firms employing at least ten employees in 1990 are grouped into 20 equally-sized bins. Actual average normalised native growth for the industry is shown as a horizontal line. Data source: *Betriebshistorikpanel*.

However, the patterns presented in Figure 1 are not necessarily representative of all industries and all time periods. In this paper, I formally test for tipping points in the composition of firms in Germany over the period 1975–2010, which I divide into five-year subperiods. Formally, the test for the presence of tipping points takes the form of a test for a downward intercept shift at some value of the base-year immigrant share in expected native workforce growth net of immigrant workforce growth. I find only limited evidence of tipping points in the composition of firms. The evidence is clearest in periods where Germany experienced relatively large inflows of immigrants and for firms operating in lower-paying, disproportionately low-skill industrial sectors. For example, for the period 1990–1995, when Germany experienced large inflows of immigrants from the former USSR and Yugoslavia, I find evidence of tipping points in six sectors out of 15, while across years I find evidence of tipping points in multiple years for predominantly lower-skill sectors including Manufacturing, Hotels and restaurants, and Transport, storage and communication.

Testing for the presence of tipping points presents econometric challenges, since the location of the tipping point is unknown to the researcher. Tests for tipping points in other settings have typically followed the method proposed by Card et al. (2008, 2011), which treats finding the location of the tipping point and testing for a discontinuity at the tipping point as separate problems. However, such a method may not be appropriate if the discontinuity is small relative to sampling variation, which will be the case if there is in fact no discontinuity at all. I therefore simultaneously identify the location of the tipping point and the size of the discontinuity via nonlinear least squares and use the methods proposed by Andrews et al. (2019, 2021) for conducting inference on the size of a discontinuity when the location of the discontinuity is unknown. Since the inference methods developed by Andrews et al. (2021) have not yet been widely used in applications, I detail how these are applied in my setting in an online appendix.

The methods developed by Andrews et al. (2021), like those proposed by Card et al. (2008), assume that the location of the tipping point is common to the group of firms for which one is testing for a tipping point. However, if there are firm-specific amenities that matter differently to natives and immigrants, the location of the tipping point might be specific to each firm (Banzhaf and Walsh, 2013; Caetano and Maheshri, 2017). In my setting, the importance of amenities is confirmed by the fact that evidence of tipping points is clearest when grouping firms by sector, rather than by labour market, consistent with sector and industry accounting for a much larger share of the variance in firm-level amenities than geography does (Sorkin, 2018).

I therefore consider numerous robustness checks where I group firms by different proxies for unobserved amenities that might be differentially valued by natives and immigrants. This is to ensure that I am not missing true tipping points by inappropriately partitioning firms into groups with a high within-group variance in firm-level amenities. In particular, I consider grouping firms by firm fixed effects estimated in an individual wage regression, consistent with recent evidence on the role of firm-level amenities in driving variation in the firm-specific component of wages (Sorkin, 2018), as well as by narrowly defined industry, and repeat the test for tipping dynamics. The general pattern of evidence remains the same. Tipping points are only identified in a subset of firms, corresponding to 15–20 per cent of groups of firms considered. The evidence is again strongest in periods of high immigration and in particular 1990–1995. I also consider only smaller firms, where interpersonal interactions between coworkers are likely strongest and find, if anything, slightly stronger evidence of tipping points.

Evidence of tipping points in the composition of neighbourhoods is robust (Card et al., 2008, 2011; Aldén et al., 2015). In contrast, while cross-sectional segregation by ethnicity or race has been widely documented in the labour market (Andersson et al., 2014; Åslund and Skans, 2010; Glitz, 2014; Hellerstein and Neumark, 2008; Higgs, 1977), the only formal test of tipping points in the labour market is Pan (2015), who finds clear evidence of tipping points in the gender composition of occupations in the US.[2] The first contribution of the paper is therefore to formally test for tipping points in a setting, namely the labour market, and specifically in the composition of firms, where the dynamics of segregation have been under-studied relative to residential segregation.

The second contribution of the paper is to an emerging literature on firm hiring of immigrants. It has been shown that firms with certain observable characteristics, namely larger firms and firms founded by immigrants, are more likely than other firms to hire immigrants (Brinatti and Morales, 2021; Kerr and Kerr, 2021). However, there also appears to be a firm life-cycle in the hiring of minorities, with firms tending to become more diverse as they age (Miller and Schmutte, 2021; see also Lepage, 2021). Here I consider how the contemporaneous immigrant share might matter *per se* for the subsequent hiring and retention of immigrants and natives. The formal test of tipping points complements recent findings on immigrant hiring by showing that there is some, albeit limited, evidence of a systematic pattern of firms switching over time from being mixed or low-immigrant to being high-immigrant firms in periods when immigrant inflows are sufficiently large,

The third contribution of the paper is to our understanding of the mechanisms underlying workplace segregation. Tipping points in firm composition are a necessary condition for the existence of strong preference spillovers. I find only limited evidence of tipping points, which suggests that explanations of workplace segregation that build on preference spillovers (e.g. building on Goldin, 2014b) are on the whole not as important for explaining observed workplace segregation as the use of referrals in hiring (Miller and Schmutte, 2021) or differences in manager hiring across nativity (Åslund et al., 2014; Lepage, 2021; Kerr and Kerr, 2021). This is similar to the evolving view in the literature on residential segregation, where the older consensus that preference spillovers are a major explanation of observed segregation (Becker and Murphy, 2000; Card et al., 2008; Schelling, 1971, 1978) has recently been challenged by theoretical (Banzhaf and Walsh, 2013) and empirical (Caetano and Maheshri, 2021) arguments that unobserved neighbourhood amenities dwarf preference-driven endogenous feedback mechanisms in explaining observed segregation. However, I present new evidence on the evolution of segregation within firm cohorts over time and show that aggregate segregation increases within a cohort in years of high immigration. This pattern suggests that preference spillovers and the feedback loops they create between past and future changes to a firm's composition may nevertheless be relev-

[2] In unpublished work, Zheng (2014) finds some evidence of tipping points in the racial composition of firms in the US, although she does not study the nativity of workers, nor are her methods robust to the null of no tipping point being true.

ant for understanding changes in segregation when firms face large shocks to the relative supplies of different types of workers.

The paper is structured as follows. In the following section I outline a model of workplace segregation and show how preference spillovers lead to discontinuities in native workforce growth. In Section 3 I present the empirical implications of the model and discuss different tests for the existence of tipping points, before presenting the data to be used in Section 4. In Section 5 I present my results, focusing on tests for the presence of tipping points in the composition of firms and exploring the robustness of the results. In Section 6 I interpret my results in relation to previous research and when compared to general trends in segregation across firms over time. Section 7 concludes.

2 Theoretical framework

2.1 A model of tipping

In this section I briefly adapt the model of Card et al. (2008, 2011) of neighbourhood composition in the presence of social interactions to segregation in the labour market. This stylised model will serve to guide the empirical analysis. The model is static and partial equilibrium. A representative, nondiscriminating firm hires two types of workers, immigrants and natives, denoted $j \in \{I, N\}$, which it treats as perfectly substitutable in production. The firm's size is taken as given, so the total workforce is normalised to equal one. The supply of workers of each type to the firm is a primitive of the model. To hire a given quantity n_j of type j, a firm needs to pay a wage $\omega^j(n_j, s)$. Crucially, the wage a firm needs to pay to hire depends not only on the quantity of workers of type j it wishes to hire, n_j, but also on the share of immigrants in the firm, s.

The partial derivatives $\partial\omega^j(n_j, s)/\partial n_j$ are assumed to be weakly positive, that is, for a constant immigrant share, the firm needs to raise wages to hire more workers of a given type. The partial derivative $\partial\omega^j(n_j, s)/\partial s$ represents the social interaction effects. In particular, similar to Card et al. (2008), I assume that $\partial\omega^N(n_N, s)/\partial s > 0$ for s greater than some threshold;[3] that is, when the immigrant share in the firm is large, the firm needs to pay a higher wage to hire a given quantity of natives.

Under the normalisation that the total workforce is one, we have $n_N = 1 - s$, and the derivative of $\omega^N(1 - s, s)$ with respect to the migrant share will be

$$\frac{d\omega^N}{ds} = -\frac{\partial\omega^N}{\partial n_N} + \frac{\partial\omega^N}{\partial s}. \tag{1}$$

[3]Note I am not assuming a discontinuity in $\partial\omega^N(n_N, s)/\partial s$; the partial derivative may vary smoothly through the threshold in s, or it may be positive for all $s \geq 0$.

Under the previous assumptions, the first term will be negative, while the second term will be positive when s is above some threshold. A tipping point in the composition of the firm's workforce can be observed if one assumes that the social interaction effect dominates, i.e. $d\omega^N/ds > 0$, at high levels of s but not at low levels of s (Card et al., 2008). The wage schedule for natives is therefore downward sloping in the quantity of natives to be hired $n_N = 1 - s$ for low levels of n_N; the reduction in s entailed by the increase in n_N increases the attractiveness of the firm sufficiently to attract more native workers, even at a lower wage. The wage schedule only becomes upward-sloping as n_N rises and the immigrant share s falls below a certain threshold. I also assume for simplicity that $d\omega^I/ds > 0$ for all $s \in (0, 1)$, that is, that the wage schedule for immigrants is upward-sloping in the quantity of immigrants to be hired for all values of n_I.[4]

There are multiple ways one could interpret the social interaction effects captured by the assumption that $\partial\omega^N(n_N, s)/\partial s > 0$. The simplest way, consistent with the original model of Card et al. (2008) and the tradition of social interactions models going back to Schelling (1971), is to interpret this as a consumption externality. Natives experience disutility from working with immigrants, so the marginal native worker will become unwilling to work at the firm if the immigrant share increases. The source of this disutility could be a simple distaste or discomfort experienced by individual natives when working with immigrants. In 2017, only 37 per cent of Germans stated they would be "totally comfortable" having an immigrant as a work colleague, similar to the proportion (36 per cent) stating that they would be totally comfortable having an immigrant as a neighbour (European Commission, 2018).[5] Alternatively, the disutility could arise from dynamic considerations, if natives believe that working with immigrants will harm their future job-finding prospects and earnings. Such beliefs could arise if immigrants are not a good source of referrals or information about job openings, or if an inflow of immigrants into a firm is a signal that the firm has experienced a negative productivity shock, as in the pollution model of Goldin (2014b).[6]

[4]There is therefore an asymmetry in the strength of the social interaction effects between immigrants and natives that drives an asymmetry in the shape of the inverse supply curves of migrants and natives. This asymmetry is also present in the model of neighbourhood composition of Card et al. (2008). The empirical predictions of the model can still be derived when social interactions cause immigrant inverse supply to be downward sloping in n_I for low values of s; what is strictly necessary however is that the inverse supply curve of immigrants be flatter than the inverse supply curve of natives, i.e. $d^2\omega^I/ds^2 < d^2\omega^N/ds^2$, for all $s \in (0, 1)$.

[5]The other options were "somewhat comfortable", "somewhat uncomfortable", "totally uncomfortable", or "don't know". Across the EU, the share "totally comfortable" was 43 per cent for neighbours and 44 per cent for colleagues.

[6]Alternatively, one could interpret the social interaction effect as a productivity externality, reinterpreting n_N as the effective supply of natives. Under this interpretation, an increase in the immigrant share lowers the productivity of natives; to keep a constant effective supply of native workers, the firm must raise the wage offered to hire more natives. This interpretation is consistent with recent evidence on negative productivity spillovers between immigrants and natives in certain firms (Glover et al., 2017). However, productivity spillovers would complicate the derivation of Equation (1), since now $n_I \neq s$.

However one interprets the social interaction effect, it is worth noting that the assumption that $\partial\omega^N(n_N, s)/\partial s > 0$ for s above some threshold is consistent with heterogeneous underlying individual preferences. If all natives dislike working with immigrants, then $\partial\omega^N(n_N, s)/\partial s > 0$ for all s. If, consistent with survey evidence (European Commission, 2018), some natives are indifferent, or even positively inclined towards working with low levels of immigrants, then it may be the case that $\partial\omega^N(n_N, s)/\partial s \leq 0$ for low values of s. The only constraint on the underlying pattern of heterogeneity in native preferences is that the number of natives who for a given wage would prefer to take their outside option rather than work at the firm is increasing in s for s sufficiently large.

At an integrated equilibrium, where both types of workers are employed at the firm, the wages paid to both types of workers must be equal, since the firm is assumed to be non-discriminating. Again under the normalisation that the total workforce is one, an integrated equilibrium therefore requires that

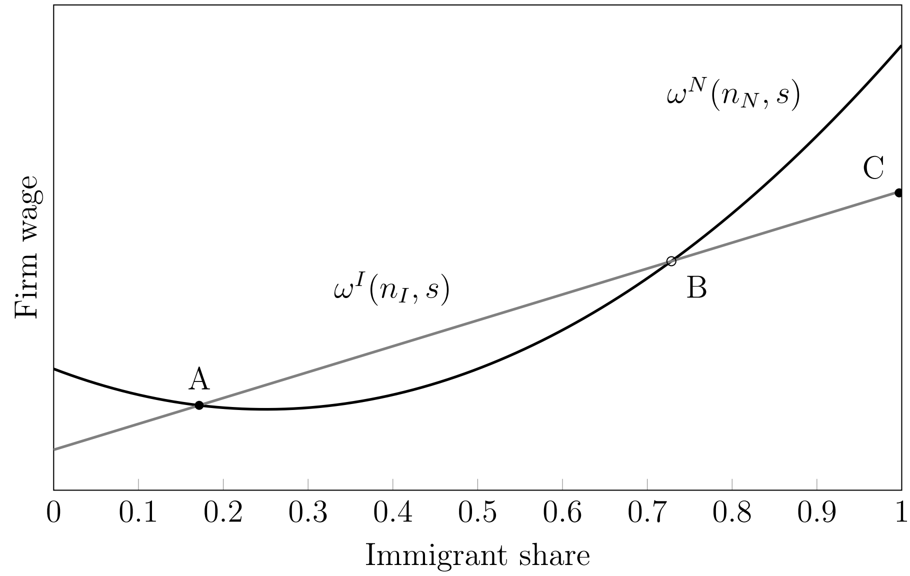
$$\omega^N(1 - s, s) = \omega^I(s, s). \tag{2}$$

The inverse supply curves of immigrants and natives are plotted in Figure 2. As $s = n_I = 1 - n_N$, the supply of immigrants increases moving to the right on the x-axis, while the supply of natives increases moving to the left on the x-axis. As the inverse supply curves are drawn, there are two integrated equilibria (A and B) and one fully segregated equilibrium (C). Equilibrium A is stable in the sense that a small increase in the firm's minority share raises the wage that must be paid to immigrants above the wage paid to natives, so the firm hires natives until it returns to the equilibrium at A. The same remark holds *mutatis mutandis* for a decrease in the minority share at A or at C. Equilibrium B is, however, unstable. After a small increase in the immigrant share from B, the wage demanded by natives is greater than the wage demanded by immigrants, the firm will replace natives with immigrants until it reaches the equilibrium at C.

In Figure 3 I plot what happens as the supply of immigrant workers to the firm increases exogenously, say, as a result of an inflow of immigrants to the local labour market where the firm is located. Suppose the firm is initially in equilibrium at E_1. An exogenous increase in the supply of immigrants shifts the immigrant inverse supply curve downward. The equilibrium moves to the right, eventually reaching the point of tangency E_2, which is stable with respect to decreases in the immigrant share, but unstable with respect to increases. If there are any further increases in the supply of immigrant workers, no integrated equilibrium will exist, the only equilibrium will involve the firm hiring only immigrants, as at point E_3. Traditional social interaction models such as Schelling (1971, 1978), Becker and Murphy (2000), or Banzhaf and Walsh (2013) would

Furthermore, the empirical implications of the model do not depend on whether the social interaction effect captures a consumption or a productivity externality, so I do not entertain this idea further here.

Figure 2: Immigrant and native inverse labour supply



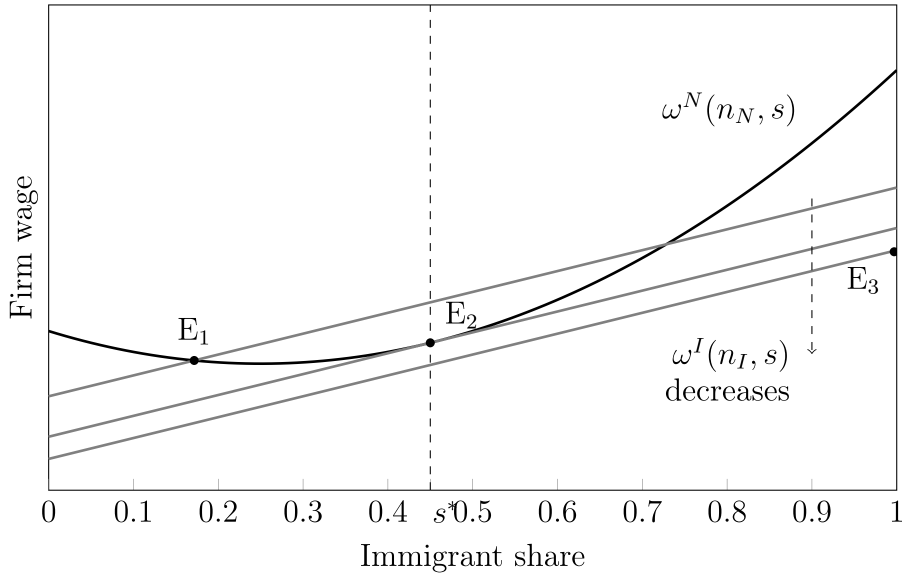
Notes: Immigrant and Native inverse labour supply to the firm with three equilibria. A and C are stable, B is unstable.

identify the unstable equilibrium B in Figure 2 as a tipping point. Here, however, I follow Card et al. (2008) in defining the tipping point as the maximum possible immigrant share in an integrated equilibrium. In Figure 3, this is the immigrant share s^*, associated with the equilibrium E_2.[7]

Two caveats are worth noting with this model. First, it does not account for the distribution of immigrants across firms, only the composition of a single firm. I implicitly assume that the natives who leave the firm after the tipping point is exceeded would either prefer to be unemployed than keep working in a high-immigrant-share firm, or are able to find jobs in other firms that have not faced a similar supply shock. Second, social interaction models are typically thought to lead to an inefficiently high degree of segregation across neighbourhoods, because agents cannot coordinate on where to locate (Becker and Murphy, 2000). The model presented here, by only considering a single representative firm, is silent about the potential welfare consequences of such social preferences. It has traditionally been argued that firms, by internalising any spillovers across workers arising from their hiring decisions, choose a socially optimal degree of segregation (Becker and Murphy, 2000). However, these arguments do not account for the possibility that workplace segregation could be dynamically inefficient, if, for example, it keeps immigrants

[7] Card-style tipping points, i.e. the largest minority share possible in a stable equilibrium, are sometimes referred to as bifurcation points (Caetano and Maheshri, 2017), to differentiated them from Schelling-style tipping points, identified as an unstable equilibrium. Bifurcation points imply that tipping is "one-sided"—firms in an integrated equilibrium might tip to being all-migrant, but not all-native—whereas Schelling-style tipping points can lead firms to tip from an integrated equilibrium to being either all-migrant or all-native (Card et al., 2011). I follow the practice of Card et al. (2008) and Pan (2015) in referring to such bifurcation points as tipping points.

Figure 3: Effect of increasing supply of immigrant labour



Notes: Increasing supply of immigrant workers shifts their relative supply outwards, decreasing the wage demanded for any value of s. The equilibrium immigrant share starts at E_1 and shifts right as the inverse supply of migrants increases. The equilibrium E_2 is the maximum integrated equilibrium, the associated migrant share is s^*. If the supply of immigrant workers increases further, the firm will jump to the segregated equilibrium E_3, hiring only immigrants.

from developing the network or the kind of experience necessary to move up the job ladder or if it prevents employers from learning the true average productivity of immigrants (Lepage, 2021).

It is also worth noting that I assume in the model that immigrants and natives are perfect substitutes. Evidence suggests that the macro elasticity of substitution between immigrants and natives of similar education and experience levels is large, but finite (Ottaviano and Peri, 2012). Since the macro elasticity captures imperfect substitutability both within and across firms (Oberfield and Raval, 2021), the micro, i.e. within-firm elasticity is likely to be even higher. For that reason I do not explicitly model the possibility of imperfect substitutability between types of workers within the firm.

2.2 Dynamic implications

While the model presented in the previous section is static, it is still possible to use it to make dynamic predictions about the composition of the representative firm's workforce.

Consider a firm whose initial static equilibrium immigrant share is $\bar{s}_0 < s^*$, where s^* is the tipping point defined previously as the immigrant share associated with the maximum possible integrated equilibrium. Suppose the firm experiences a small increase in the supply of immigrants, i.e. a fall in the wage a given quantity of immigrant labour needs to be paid, $\Delta\omega^I(n_I, s) < 0$, between period 0 and period 1.[8] There will be some $r \in (0, s^*)$ such that if $\bar{s}_0 \in [0, s^* - r)$, the firm's new equilibrium will be at $\bar{s}_1 \in (0, s^*]$, whereas if $\bar{s}_0 \in [s^* - r, s^*]$, the increase in the immigrant supply takes the firm beyond the point of tangency at E_2 in Figure 3 and the new equilibrium will be $\bar{s}_1 = 1$. As the increase in the immigrant supply $\Delta\omega^I(n_I, s)$ becomes infinitesimally small, r also approaches zero. Note that no firm can initially be at an equilibrium at $\bar{s}_0 \in (s^*, 1]$ except for at $\bar{s}_0 = 1$, where a small increase in the supply of immigrants will have no effect on the equilibrium.

Assume that the firm myopically adjusts its immigrant share in response to changes in the supply of immigrants such that the immigrant share s_t remains close to its equilibrium value. To allow for the possibility that search or other labour market frictions prevent the immigrant share from fully adjusting within a single period to a new equilibrium value as the supply of immigrants changes, I use the notation s_t to refer to the observed immigrant share at a point in time, to distinguish it from the static equilibrium at that point in time, $\bar{s}_t$. For an observed $s_0 \in [0, s^* - r)$, the observed increase in the immigrant share Δs_1 in response to the increase in the immigrant supply $\Delta\omega^I(n_I, s)$ will be small. However, for $s_0 \in [s^* - r, s^*]$, $\Delta\omega^I(n_I, s)$ will cause a large observed Δs_1, as the firm converges to the new equilibrium at $\bar{s}_1 = 1$. For firms initially at $s_0 \in (s^*, 1)$, the tipping process is already underway, and one should expect to see $\Delta s_1 > 0$ and larger the closer the firm is to s^*. There will therefore be a discontinuity in Δs_1 around the tipping point s^*. We will observe Δs_1 to be small and positive for s_0 to the left of the tipping point and large and positive for s_0 close to or beyond the tipping point.

Whilst the foregoing discussion restricts attention to the case of an increase in the immigrant supply, where the discontinuity appears clearly, the discontinuity will also exist in the case where there is a decrease in the immigrant supply. This is because once a firm has started tipping and $s_0 \in (s^*, 1]$, a small decrease in the supply of immigrants will typically not reverse the tipping process, implying that for these firms too $\Delta s_1 > 0$. The condition for tipping to continue after a decrease in the immigrant supply is for the marginal immigrant to continue to accept a lower wage than the marginal native, which is more likely to be satisfied the smaller the decrease in the immigrant supply or the further to the right of s^* the firm initially finds itself. On the other hand, for a firm that is close to tipping, but where $s_0 < s^*$, a small decrease in the immigrant supply will lead to a small decrease in the immigrant share in the firm.

Combining these observations about the effect of increases and decreases in the immigrant supply on the firm's immigrant share, one can conclude that there will be a discontinuity in the expected change in the immigrant share as a function of the base-year immigrant share:

[8] The discussion here in fact holds for an increase in the relative supply of immigrant, $\omega^N(n_N, s) - \omega^I(n_I, s)$. However, to simplify the discussion I assume the supply of natives is fixed and only the supply of immigrants varies.

$$\mathrm{E}[\Delta s_t|s_{t-1}] = \mathbf{1}(s_{t-1} < s^*)g(s_{t-1}) + \mathbf{1}(s_{t-1} \geq s^*)h(s_{t-1}) \tag{3}$$

where $\lim_{\epsilon \to 0^+} h(s^* + \epsilon) - g(s^* - \epsilon) > 0$. $h(s_{t-1}) > 0$, while the sign of $g(s_{t-1})$ will depend on whether firms more commonly face increases or decreases in the immigrant supply. The existence of a discontinuity in $\mathrm{E}[\Delta s_t|s_{t-1}]$ at the tipping point s^*, which does not depend on whether the immigrant supply is increasing or decreasing, is the key dynamic implication of the model I will test in the empirical analysis below.

3 Empirical implementation

3.1 Unit of analysis

While the model presented above predicts that tipping points might be observed in the composition of a firm's workforce, one might also expect to observe tipping dynamics in the composition of larger aggregates, such as the industry, occupation, or geographic area. Indeed Goldin (2014b) notes that the pollution model she develops to explain the dynamics of workplace composition by gender might operate at the level of firms, occupations, industries, or geographic aggregates. Historically, there is evidence in France at least of high immigrant shares in an industry being associated with low prestige of the industry (Noiriel, 1988), suggesting that tipping might occur in the composition of larger aggregates. On the other hand, if the kinds of preference spillovers underpinning the model of tipping presented above are experienced primarily in direct personal interactions in the workplace, as in the cases studied by Hjort (2014) or Glover et al. (2017), one might expect to only observe tipping in the composition of firms.[9]

Empirically, segregation across firms and across larger units of aggregation appear to be distinct phenomena. Table 1 reports the index of coworker segregation—defined by Hellerstein and Neumark (2008) as the excess probability that an immigrant has of working with other immigrants, relative to a native—for West Germany in 1985–2010. Throughout this period, an immigrant was at least 16 percentage points more likely to work with another immigrant than natives were. The index of coworker segregation can be normalised to account for differences in the distribution of immigrants and natives across larger units of aggregation, such as regions or industries. This is done by repeatedly simulating a counterfactual distribution of immigrants over firms, conditional on the observed immigrant shares in regions or industries. The average counterfactual index of coworker

[9]Note also that there is no straightforward logical relationship between tipping at, say, the industry level and at the firm level. Industry-level tipping does not imply firm-level tipping, since it could occur through the entry of high immigrant-share or the exit of high native-share firms as the industry passes the tipping point. Similarly, tipping at the firm level might only imply a reallocation of a fixed pool of workers within the industry, leaving the aggregate composition unchanged.

Table 1: Index of coworker segregation

	1985 ICS	1990 ICS	1995 ICS	2000 ICS	2005 ICS
Unconditional	0.16	0.16	0.18	0.19	0.18
Conditional on industry	0.12	0.12	0.13	0.14	0.14
Conditional on location	0.14	0.14	0.16	0.17	0.17
Conditional on location and industry	0.08	0.08	0.09	0.10	0.10
Establishments	480174	520287	545227	721179	745427

Note: Indexes of coworker segregation of Hellerstein and Neumark (2008), calculated from the *Betriebshistorikpanel* of the IAB. Includes all establishments in West Germany employing two or more workers. The conditional indexes condition on either three-digit industry (NACE Rev. 1), local labour market (Kropp and Schwengler, 2011), or both.

segregation is then subtracted from the true index, yielding what is known as an *effective* index of coworker segregation.

Conditioning the index on the distribution of workers over local labour markets and three-digit industries reduces an immigrant's excess probability of working with other immigrants to 8–10 percentage points, explaining 45–50 per cent of observed segregation, with segregation across industries appearing to contribute more to this reduction than segregation across locations. By way of comparison, Glitz (2014) finds an unconditional index of residential segregation across municipalities in Germany of 0.07, while the effective index of residential segregation, conditioning on region of residence, is 0.05. In the analysis in Section 5 I will mainly focus on testing for tipping points in the composition of firms, given that individuals will interact more intensively with their colleagues than with workers at other firms in the same industry. However, given that segregation across firms and across geographically delimited industries are theoretically and empirically distinct phenomena, in a robustness check I will also investigate the presence of tipping points in local industries, defined as the aggregation of all establishment operating in a three-digit industry in a given local labour market (defined using commuter flows, see Kropp and Schwengler, 2011).

3.2 Identifying the location of the tipping point

Any test for the existence of a tipping point in workforce composition needs to reckon with the fact that the theoretical tipping point s^* is unknown. Card et al. (2008) propose treating identifying the location of the tipping point and testing for the existence of a tipping point as separate problems and solving them sequentially. In the first step, they use a search procedure to identify a candidate tipping point. The simplest procedure they propose is a threshold regression (Hansen, 2011, 2021). In the second step, Card et al. (2008) use regression discontinuity design (RDD) techniques (Imbens and Lemieux, 2008; Lee and Lemieux, 2010) to estimate Equation (3). If the estimated discontinuity in the change in the minority share when the minority share moves beyond the candidate tipping point is negative and significant, they conclude that there is a tipping point in the composition of the units under study.

To address the possibility of specification search bias that would arise when using the same data to both identify the location of the tipping point and estimate the size of the discontinuity at the tipping point, Card et al. (2008) propose two solutions. Either the researcher can split the sample, using independent subsamples for the search and estimation steps described above, or the researcher may bootstrap the entire two-step procedure to construct standard errors for the estimated second-stage discontinuity.

The estimation and inference procedure proposed by Card et al. (2008) has been adopted, essentially unmodified, in many subsequent tests of tipping points (Aldén et al., 2015; Böhlmark and Willén, 2020; Pan, 2015; Zheng, 2014).[10] However, the approach suffers from two shortcomings. First, treating the second stage as an RDD is arguably conceptually incorrect, since there is no treatment variable whose assignment probability jumps at the threshold, other than the tautologically defined treatment "being above the tipping point". So while one may still use local polynomials to descriptively estimate a break in the outcome variable at the candidate tipping point, the standard RDD interpretation of this break as an average treatment effect does not apply.[11] Second, the inference procedures proposed by Card et al. (2008) may not be suitable in all settings, and in particular in settings where there is in fact no tipping point.

Conducting inference via sample splitting is not efficient, since only a subset of the data is used at either stage. The approach therefore relies on the availability of a large dataset, which is the case when studying tipping in firms or neighbourhoods, of which there are many, but not when studying larger aggregates, such as local industries. Conducting inference via the bootstrap, on the other hand, is feasible in smaller datasets, but its validity has been demonstrated under the assumption that the discontinuity being estimated is large relative to sampling variation (Hansen, 1996; Elliott and Müller, 2007). However, in situations where it is not obvious from simply looking at the data whether there is a tipping point or not, such an approach may lead to over-rejection of the null hypothesis of no tipping points (Andrews et al., 2021). In such settings, the researcher may conclude that there are tipping points where there are in fact none.

To address these shortcomings, rather than treating the problem as a type of RDD,

[10]These papers all work with a Card-style definition of a tipping point as a bifurcation point. Alternative methods have been developed to test for Schelling-style tipping points as unstable equilibria, see Caetano and Maheshri (2017, 2023).

[11]This criticism does not apply to the work of Böhlmark and Willén (2020), since for their main results they use tipping points identified via a Card-style procedure as discontinuities in an unrelated RDD, where the main outcome is individual-level educational attainment.

I use methods from the literature on structural breaks to identify both the location and the size of the discontinuity at the potential tipping point and use inference procedures that are robust to small effects. Both the location of the tipping point and the size of the break in the outcome are estimated via a threshold regression that takes the following general form:

$$Y_{it} = C_{it}'\beta + D_{it}'\delta\mathbf{1}\{Q_{it} > \theta\} + u_{it}. \tag{4}$$

Let the number of immigrants employed in firm i at time t be I_{it}, the number of natives be N_{it}, and the total workforce $L_{it} = I_{it} + N_{it}$. Following Pan (2015), the dependent variable Y_{it} is defined as the five-year change in the native workforce, normalised by the base-year workforce, minus the normalised five-year change in the immigrant workforce: $Y_{it} = (N_{it+5} - N_{it})/L_{it} - (I_{it+5} - I_{it})/L_{it}$.[12] The change in immigrant demand is therefore a proxy for changes in total workforce demand, which are netted out in this formulation (Pan, 2015). The vector of control variables C_{it} includes a polynomial function in the base-year immigrant share and other base-year controls depending on whether the unit of observation is the firm or industry. Q_{it} is the base year immigrant share and θ is the tipping point. The set of variables D_{it} is the subset of C_{it} whose effect on Y_{it} varies when the base-year immigrant share passes the tipping point. In my specifications D_{it} only includes a constant; in this case, the parameter δ measures the key discontinuity. We conclude that there is a tipping point if δ is negative and significant. The estimation Equation (4) is the empirical counterpart of Equation (3).

Equation (4) is nonlinear in the parameter vector $(\beta', \delta', \theta)'$, and is estimated by non-linear least squares (NLS). The location of the tipping point, θ, and the size of the discontinuity at the tipping point, δ, are therefore estimated simultaneously. The difference between this approach and that of Card et al. (2008) bears emphasising. They estimate the location of the candidate tipping point s^* from a simple threshold regression where $C_{it} = D_{it} = \iota$, a constant, and then estimate $\delta(s^*)$ from a follow-up OLS regression of Equation (4), which they characterise as an RDD, where they set $\theta = s^*$ and C_{it} includes higher-order polynomial terms and other controls.

While Equation (4) can be estimated by NLS, the parameters are not asymptotically normally distributed, since θ is not identified when $\delta = 0$ (Hansen, 2021). Hansen (1996) has shown that a bootstrap procedure will yield correct p-values for the test that $\delta = 0$, and Card et al. (2008) appeal to this result when justifying the use of the bootstrap to construct standard errors for $\delta(s^*)$ in their two-step procedure. However, the validity of the bootstrap procedure in the threshold regression setting is shown under a restrictive

[12] Other tests of tipping points, such as Card et al. (2008) or Pan (2015), study ten-year changes. In part this is due to data constraints, since those papers use decennial census data. Since the costs of changing workplace are arguably lower than changing residence, workplace tipping dynamics will appear on a shorter time scale. Focusing on ten-year changes would also lead to greater selection into the sample via firm exit, which is potentially correlated with the base-year immigrant share.

set of assumptions (Andrews et al., 2021). As a result, Andrews et al. (2021) propose an alternative procedure for constructing standard errors for δ when estimating a threshold regression. In particular, their procedure is robust to (i) the true threshold effect δ being small relative to sampling variation; and (ii) the model (4) being misspecified, which is likely if Equation (4) is only a parsimonious approximation of the true conditional expectation of Y_{it}. I will therefore use the so-called "hybrid" standard errors proposed by Andrews et al. (2019, 2021) when conducting inference on δ. These standard errors have been shown both theoretically and in simulations to have good coverage properties both when the truth is $\delta = 0$ and when $\delta \neq 0$. The interested reader is referred to Andrews et al. (2019, 2021) for full details on the construction of these standard errors.[13]

3.3 Variation in the location of the tipping point

The location of any tipping point, $\theta = s^*$, will depend on the shape of the inverse supply curves of immigrant and native workers. Various factors will affect the shape of the inverse supply functions, two of which bear emphasising here. The first factor is heterogeneity across locations in individual tastes within the pool of workers the firm might potentially hire, and, in particular, heterogeneity in the strength of native distaste for immigrants, measured by the partial derivative $\partial \omega^j(n_j, s)/\partial s$, which might vary with historical exposure to immigration. If the value of the partial derivatives of the inverse supply functions is the same across locations, then the tipping point will also be the same for different locations.

The second important factor that will affect the location of the tipping point are firm-specific amenities differentially valued by natives and immigrants. Such amenities may make some firms more attractive to natives than others, for a given wage and immigrant share, potentially altering the shape of the native inverse supply curve. If such amenities vary substantially across firms, the location of the tipping point will also vary across firms.

The procedure presented in Section 3.2 assumes that the location of the tipping point is common to at least some subset of firms, i.e. that it is possible to group firms by some combination of location and a proxy for non-wage amenities valued by natives before testing for a common tipping point for a given grouping of firms. Both Card et al. (2008) and Aldén et al. (2015) assume different tipping points for different residential markets (metropolitan areas), while Pan (2015) assumes the location of tipping points in labour markets varies by region-occupation type (white/blue collar) cell. However, the import-

[13] Andrews et al. (2021) develop their procedure in the case where $D_i = C_i$. Since in my setting D_i is typically a constant while C_i includes a polynomial in the base-year immigrant share, and since the method of Andrews et al. (2021) has not yet to my knowledge been used in applications, I present in detail the changes that are necessary to implement their method and construct standard errors when D_i is a subset of C_i in Appendix B, available online. These details may be of use to researchers interested in implementing the procedure of Andrews et al. (2021) in other settings.

ance of heterogeneous neighbourhood amenities in the dynamics of residential segregation has been highlighted theoretically by Banzhaf and Walsh (2013) and demonstrated empirically in the case of school segregation by Caetano and Maheshri (2017);[14] there is no reason to suppose there is less heterogeneity in firm amenities than in neighbourhood amenities. If firms are grouped in a way that does not adequately capture underlying heterogeneity in amenities or preferences, we might fail to find evidence of tipping points even though they might actually exist in practice.

In Table 2, I report naive descriptive evidence on the presence of tipping-like dynamics in alternative groupings of firms. In each grouping of firms, e.g. two-digit industry, I calculate the average of normalised five-year native workforce growth over firms with an above-average initial immigrant share and firms with a below-average initial immigrant share. I characterise a group of firms as following a tipping-like dynamic if the average normalised native workforce growth is lower for firms with an above-average immigrant share. I consider grouping firms by local labour markets, two-digit industries, or the intersection of labour market and industry.

Table 2: Naive evidence of tipping-like dynamics

	1985	1990	1995	2000	2005
	(1)	(2)	(3)	(4)	(5)
Local labour markets	0.10	0.64	0.30	0.04	0.05
Two-digit industries	0.41	0.50	0.38	0.19	0.28
Industry-labour markets	0.50	0.49	0.32	0.39	0.37

Note: Share of different types of cells where average normalised native workforce growth is lower for firms with an above-average initial immigrant share. Source: *Betriebshistorikpanel* of the IAB. Calculated using establishments in West Germany employing ten or more workers.

Firms display behaviour consistent with tipping-like dynamics when grouped by either industry or by labour market, particularly in the earlier part of the sample, pre-2000, where the fraction of either industries or labour markets where tipping-like dynamics are observed is around 0.5. If workers' nativity were irrelevant to firms' hiring decisions, firm-level deviations from the average immigrant share would be transient and we would observe mean-reversion in the firm's immigrant share over time. The probability that average normalised native workforce growth is lower for firms with above-average immigrant share would therefore approach zero for sufficiently large groupings of firms. The descriptive evidence in Table 2 therefore already provides initial support for the existence

[14] Caetano and Maheshri (2017) propose a method for testing for the presence of school-specific tipping points in school composition given heterogeneous school amenities. Extending their approach to the case of firms, which would require a credible instrument for the immigrant share in the firm, is beyond the scope of the present work.

of tipping points in the composition of firms, even when firms are grouped by relatively broad aggregates.

However, the evidence in Table 2 does not clearly point to labour markets or industries as a better proxy for the underlying factors that determine the location of the tipping point; indeed the evidence for tipping is arguably strongest when both are used to group firms. There is, however, a cost to grouping the data in overly small cells. When identifying tipping points using the threshold regression described in Section 3.2, the location of the tipping point is only identified by observations close to the threshold, so the sample needs to be relatively large. Hansen (2021) suggests $n = 500$ as a reasonable minimum size, however this is just a rule of thumb, not a theory.

In the main analysis I will therefore use single-letter industrial sector codes and labour market regions, equivalent to commuting zones in the USA, as possible groupings of firms with a common tipping point, to ensure that there are enough observations within each cell to identify a tipping point, if one exists. I will estimate Equation (4) separately for each proposed grouping and for each base year, since the descriptive evidence suggests tipping dynamics might be observed in some years but not others. I will then consider finer alternate groupings, including the interaction of broad regional and industrial groups, in a set of robustness checks. I will also consider grouping firms by estimated wage fixed effect from an AKM-type regression, as an alternative proxy for firm-level amenities; Sorkin (2018) has found that 70 per cent of the variance of the firm component of wages, estimated as firm fixed effects, reflect compensating differentials for firm-level amenities.

4 Data

The data used to test for the presence of tipping points in the German labour market come from the Institute for Employment Research of the German Federal Employment Agency (IAB). I use the Establishment History Panel (*Betriebshistorikpanel*, BHP), a fifty per cent sample of all establishments making social security contributions for at least one employee between 1975 and 2019.[15] An establishment covers all production sites belonging to the same firm, located within the same municipality, and operating within the same three-digit sector. I follow standard practice when working with the BHP in indiscriminately referring to establishments as firms or establishments.

The sampling frame of the BHP includes all firms making social security contributions in West Germany since 1975, and all such firms in East Germany since 1993. Establishment variables are measured on 30 June of a given year; an establishment may drop-out of the data if it employs no individuals subject to social security on 30 June of the ob-

[15] Specifically, I use version 2 of the 1975–2019 edition of the BHP. For details on this dataset, see Ganzer et al. (2021).

servation year and reappears in a later year. Employees exclude civil servants and the self-employed, immigrant status is defined in the data by citizenship, rather than country of birth. I restrict my attention to the period 1975–2010 and in the main analysis separately analyse changes over each of the seven five-year periods in the dataset, starting from 1975–1980. This allows me to investigate potential differences in tipping dynamics as immigrant flows and macroeconomic conditions change over time. I also limit myself to West Germany (excluding Berlin) since East Germany is not covered through the whole period and a large majority of Germany's immigrants live and work in the old West Germany.

I test for tipping dynamics in both the composition of firms and of local industries. When studying firms, I impose the supplementary restriction that firms employ at least 10 workers in the base year. I do this since (i) the immigrant share variable is not continuous when there are few employees and has mass points around values such as 0.25, 0.33, or 0.5, while the theory developed in Section 2 assumes the immigrant share is continuously distributed; (ii) the immigrant share can change dramatically over time when there are only a few workers, creating artificial discontinuities in Y_{it} around the values of the base year immigrant share where there are mass points; and (iii) small firms are more likely to enter or exit over a five-year period, potentially creating sample selection issues. When studying local industries, I similarly impose the restriction that the industry be constituted of at least ten firms, employing at least 30 workers between them. I further exclude both industries and firms where either the normalised native or immigrant workforce growth exceeds 300 per cent over five years, since the theory in Section 2 assumes the firms size is constant.

Aggregate summary statistics, using all BHP firms in West Germany, are presented in Panel A of Table 3. Averages over the firms included in my sample are in Panel B, while averages over the included local industries are in Panel C. The size restrictions imposed mean that the sample of firms cover around 62–65 per cent of total employment subject to social security in West Germany, while the sample of local industries covers around 84–92 per cent of employment. The average immigrant share in the firms and local industries studied is a little lower than in the full set of firms, implying, via Bayes's rule, that the sample of firms covers around 50 per cent of employed immigrants in Germany, while the sample of local industries covers around 65 per cent of employed immigrants.

At both levels of observation, the average immigrant share falls during 1975–1985, increases in 1985–1995, and falls again somewhat thereafter, similar to patterns of net migration to Germany over the time period. Given that, according to the theory in Section 2, tipping dynamics might be observed when there is an increase in the relative supply of immigrants facing a firm, this suggests that tipping dynamics are more likely to be observed in the period 1985–1995. While average normalised native and immigrant workforce growth are broadly correlated, there are periods, in particular 1990–1995 when high

Table 3: Summary statistics

	(1) 1975	(2) 1980	(3) 1985	(4) 1990	(5) 1995	(6) 2000	(7) 2005
A: Aggregate Statistics							
Immigrant share	9.5	9.2	7.4	8.0	7.7	7.0	7.0
Employment growth	4.0	-1.9	8.1	-3.2	12.6	-1.5	5.9
Native growth	4.1	-0.08	7.0	-4.2	12.4	-1.0	5.3
Immigrant growth	-0.05	-1.8	1.1	1.0	0.2	-0.4	0.5
B: Firm Statistics							
Share of employment	62.3	62.1	62.1	62.5	48.8	50.4	52.3
Immigrant share	7.1	6.6	5.6	6.2	8.1	6.6	6.6
Employment growth	4.8	-4.1	7.6	1.2	15.7	-1.3	5.1
Native growth	4.7	-2.9	6.4	-1.0	16.0	-1.0	4.4
Immigrant growth	0.1	-1.2	1.2	2.2	-0.4	-0.3	0.7
Firms	96359	106319	105202	115477	117904	158178	167454
C: Industry Statistics							>
Share of employment	81.0	82.3	82.8	83.9	68.3	73.9	75.6
Immigrant share	7.0	6.6	5.3	6.1	8.0	6.5	6.3
Employment growth	15.8	4.3	17.8	8.3	46.9	6.7	11.1
Native growth	15.2	5.3	15.9	5.7	44.8	6.3	9.8
Immigrant growth	0.6	-1.0	1.9	2.6	2.0	0.4	1.3
Local industries	5725	6093	6308	6746	6802	7894	7935

Note: Panel A reports aggregate statistics for all of West Germany using the BHP of the IAB. Panel B reports averages for the included firms; Panel C reports averages for the included local industries (three-digit industries by local labour markets). Growth rates are expressed in percentage terms for the five-year period starting in the base year defined for each column. Immigrant growth and native growth are normalised by total base-year employment.

immigrant inflows, in this case linked to wars in ex-Yugoslavia, coincided with protracted recessions, leading immigrant employment to grow on average even as total employment contracted.[16]

5 Results

5.1 Tipping points in firms

To test for the presence of tipping points in firms, I estimate Equation (4) separately for different groups of firms. The dependent variable is modelled as a third-order polynomial in the base-year immigrant share, with an intercept shift at the tipping point. I include the log of the median wage of a native in the firm, the low-skilled workforce share, and the firm's share of total employment in the local industry as additional controls that capture firm-specific amenities that might otherwise affect workforce growth. I present the results in Table 4. In each panel I consider a separate grouping of firms. In Panel A I group firms by single-letter industrial sector (NACE revision 1), in Panel B I group firms by regional labour markets (Kropp and Schwengler, 2011), while in Panel C I group firms by the intersection of regional labour market and an indicator for being in a low-skill industry.[17] In each panel and for each year I report summary statistics on the location of the discontinuity identified by the threshold regressions, the average estimated discontinuity, the share of cells (e.g. sectors) where the estimated discontinuity is negative and significant, as well as the median lower and upper bounds across cells of a 95 per cent confidence interval for the estimated discontinuity, the number of cells for which a regression is estimated, and the median number of observations for each regression.

The estimated discontinuities when grouping firms by sector are reported in Panel A. Evidence of tipping points in the composition of firms varies across years. On average, the threshold regression identifies a discontinuity in normalised native workforce growth at base-year immigrant shares of around 25–35 per cent of the workforce, while the average estimated discontinuity is clearly negative in some years—1980, 1990, 1995—but not all. In particular, the share of sectors in which a tipping point, i.e. a negative and statistically significant discontinuity in normalised native workforce growth, is identified varies from a low of zero in 2000 to a high of 0.4, or six sectors out of 15, in 1990, similar to the variation in the naive evidence of tipping-like dynamics presented in Table 2. Pooling all years, I find a negative and significant discontinuity in 15 per cent of sector-years.

It is important to stress that the procedure used here to construct the standard errors has been shown theoretically and in simulations to have correct coverage rates, even

[16] Germans migrating from the old East Germany to the West are classified as Germans, not migrants.

[17] Low-skill industries include Agriculture, Hunting and forestry, Fishing, Mining and quarrying, Manufacturing, Construction, and Hotels and restaurants.

Table 4: Tipping points in the composition of firms

	(1) 1975	(2) 1980	(3) 1985	(4) 1990	(5) 1995	(6) 2000	(7) 2005
A: Industrial sector							
Tipping point	38.6	35.6	27.6	30.2	38.7	23.5	38.5
	(28.3)	(31.5)	(31.9)	(32.0)	(33.9)	(24.1)	(33.7)
Discontinuity ($\hat{\delta}$)	12.7	-30.8	-1.0	-20.4	-34.2	10.6	18.7
	(58.5)	(54.2)	(39.4)	(84.6)	(71.9)	(47.0)	(51.9)
$\hat{\delta} < 0$ and p-val. < 0.05	0.13	0.27	0.13	0.40	0.07	0	0.13
Median LB, 95% CI	-21.46	-40.13	-29.20	-24.65	-26.14	-4.88	-17.21
Median UB, 95% CI	47.44	15.97	32.93	6.46	16.74	33.96	32.78
Cells	15	15	15	15	14	15	15
Median obs.	3859	4372	4474	4767	5248	7697	8481
B: Regional labour market							
Tipping point	27.0	33.6	19.8	26.3	28.5	33.2	27.9
	(22.2)	(26.8)	(20.7)	(23.2)	(23.1)	(29.9)	(24.1)
Discontinuity ($\hat{\delta}$)	11.6	-2.6	-3.9	-0.04	11.5	9.7	7.7
	(58.6)	(67.3)	(57.6)	(66.2)	(80.0)	(72.5)	(77.7)
$\hat{\delta} < 0$ and p-val. < 0.05	0.08	0.15	0.13	0.18	0.21	0.13	0.18
Median LB, 95% CI	-35.98	-39.30	-28.81	-29.31	-28.14	-34.90	-24.82
Median UB, 95% CI	50.07	35.23	27.71	29.94	44.92	30.13	28.56
Cells	39	39	39	39	39	39	39
Median obs.	1051	1209	1258	1393	1456	1879	2015
C: Region-sector type							
Tipping point	26.8	26.4	19.4	25.2	29.7	30.0	33.0
	(22.0)	(23.3)	(17.9)	(22.6)	(25.7)	(26.3)	(26.9)
Discontinuity ($\hat{\delta}$)	-46.7	51.3	5.5	22.0	11.2	3.5	7.5
	(475.9)	(451.6)	(69.9)	(125.1)	(100.9)	(100.3)	(63.4)
$\hat{\delta} < 0$ and p-val. < 0.05	0.11	0.12	0.13	0.10	0.14	0.09	0.17
Median LB, 95% CI	-47.57	-39.18	-45.53	-24.75	-48.48	-32.22	-32.65
Median UB, 95% CI	56.08	51.40	51.15	44.97	52.60	43.42	40.91
Cells	75	76	78	78	78	78	78
Median obs.	575	628	585	657	683	887	933

Note: Summary statistics on for a set of threshold regressions. In Panel A each regression uses firms from a given single-letter industrial sector (NACE Rev. 1), in Panel B a regression uses firms from a regional labour market (Kropp and Schwengler, 2011), in Panel C a regression uses firms of a given skill level (high or low) in a given labour market. Inference is conducted using the methods proposed by Andrews et al. (2021).

when the true discontinuity in the threshold model is in fact zero (Andrews et al., 2019, 2021). As a result, if there were no tipping points in the composition of firms in any industry, we should find a significant discontinuity in around 5 per cent of industry cells. Finding a negative and significant discontinuity in normalised native workforce growth in 15 per cent of industry-years, should, therefore, be interpreted *a priori* as evidence that there are probably tipping points in the composition of firms in some industries and some periods of time, but not all.

To understand in which sector-years I am more likely to identify a tipping point, and to further establish that the estimated discontinuities in Panel A of Table 4 are not the result of random chance, first consider that the years in which evidence of tipping points is strongest closely follow the periods in my sample when net immigration to Germany was positive: 1976–1981 and, in particular, 1987–1996, which saw a near-doubling of the immigrant population.[18] This pattern is consistent with the model in Section 2, which predicted that tipping might be observed in the event of an increase in the relative supply of immigrants.[19] Second, in Figure 4, I correlate various average characteristics in a sector-year with an indicator for a tipping point being identified in a sector. Here I find a clear pattern; namely, that sectors with tipping points have on average less-skilled workforces, earning lower wages. The immigrant share in a sector in the base year also appears somewhat related to the likelihood of identifying a tipping point, though not the net inflow of immigrant workers to the sector, normalised by the total base-year workforce. The importance of skill levels in whether I identify a tipping point is corroborated by looking at which sectors are the ones where tipping points are identified, reported it Table A.1. In particular, tipping points are identified in more than one year in low-skill sectors only, including Manufacturing, Hotels and restaurants, and Transport, storage and communication.

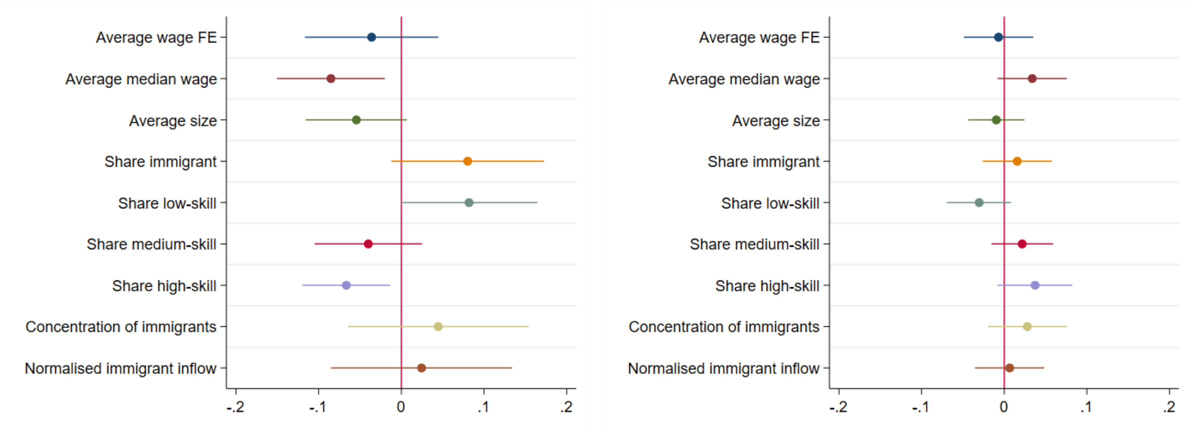
Next I consider whether there are tipping points in the composition of firms when I assume that the location of the tipping point is common to firms in the same regional labour market. Similarly to when grouping firms by sector, I identify tipping points in 15 per cent of labour market-years, however there is a less clear pattern of variation over time and over cells in where tipping points are identified. Furthermore, the average estimated discontinuity is usually positive, and is never smaller than -3.9. Compared to sectors, the correlations between labour market-year characteristics and an indicator for tipping do not suggest anything strongly predicts which labour markets will experience tipping. If anything, tipping points are more likely to be identified in labour markets with

[18] Given that immigrants typically take several years to find their first job, it is perhaps natural that the years where the largest shares of sectors display tipping points are the periods starting a couple of years after the start of each migration wave.

[19] Legislative changes also made it easier for long-term immigrants to acquire German citizenship after 2000, which could also confound patterns of tipping immediately after the reform, since immigrant status is defined in the data by citizenship, not country of birth.

Figure 4: Correlates of cell-level tipping

(a) Sector cells (b) Labour market cells



Notes: Bivariate regressions, pooling cells and years, of an indicator for tipping being observed in a cell-year on aggregate cell-year-level characteristics. Averages refer to unweighted averages across firms, shares refer to the share of workers in the cell with a given characteristic. Concentration of immigrants is the share of immigrant workers in the country employed in a given cell, normalised immigrant inflow is the five-year change in the immigrant workforce in the cell divided by the total base-year workforce. All characteristics are normalised to have mean zero and standard deviation one; the reported coefficients are the effect of a one-standard-deviation increase in a characteristic on the probability of observing a tipping point in a cell-year. N = 104 in the case of sectors, N = 273 in the case of labour markets. Robust (HC3) standard errors reported.

higher wages and skill levels, though these associations are not statistically significant; this could be because wages and skill levels are higher in larger cities, where immigrants tend to concentrate and where natives have more outside options if their firm starts hiring too many immigrants for the native's liking. The list of labour markets where I identify tipping points is reported in Table A.2.

Given that the evidence of tipping across years was stronger when grouping firms by sector than when grouping by labour market, I next consider using both to group firms. However, since grouping firms by the intersection of both variables would frequently leave me with too few observations to estimate a threshold regression, I simply group firms by the intersection of regional labour market and an indicator for being in a low-skill industry. The results are reported in Panel C. The general time-pattern of where tipping points are identified more-closely follows the pattern for regional labour markets than for industries. Overall, the evidence is slightly weaker, with tipping points identified in 12 per cent of labour market-industry type-years, although the median number of observations in a cell is relatively small, suggesting power might be an issue. Tipping points are also equally distributed over high- and low-skill cells; specifically, the probability of observing a tipping point in a high-skill labour market cell is 12.3, while it is 12.1 for a low-skill labour market cell. All in all, the results in Panel C tend to cast doubt on the existence of tipping points in geographical groupings of firms.

5.2 Robustness

5.2.1 Alternate groupings of firms

The evidence of tipping points is strongest when firms are grouped by industrial sector. This perhaps reflects that differences in amenities between firms, which are in part captured by industrial sector, are a more important determinant of variation in the location of the relevant tipping point than geographic variation in natives' preferences, which might be caused by differences in historical exposure to immigrants. However, firm-specific amenities also vary within sectors, which might lead to variation in the location of firm-specific tipping points even within sectors. To investigate this possibility further, I consider two alternative groupings of firms.

First, I consider three-digit industries, since Sorkin (2018) finds that 45 per cent of the variation in firm compensating differentials, a function of firm-level amenities, is explained by narrowly-defined industry, which suggests that most firms within sufficiently small industry cells would share a tipping point, if the tipping point exists. The results from these specifications are in Panel A of Table A.3. The evidence here for the existence of tipping points is similar to when grouping firms by coarser industrial sectors. On average, across years, I conclude that there is a tipping point in 16 per cent of industries, or around 20–30 out of approximately 145 industries; the evidence is strongest in 1980–1985, when I conclude that there are tipping points in 30 industries out of 141. However, given that there are relatively few observations in each industry-cell, as compared to the when grouping firms by sector, it may be that I lack power in some cases to detect tipping points, even though they might exist.

Second, I consider grouping firms directly by a measure of firm-level amenities that might be differentially valued by natives and immigrants. I divide firms into ventiles of the firm fixed effect from an individual wage regression.[20] Variation in firm wage fixed effects has been shown to be largely driven by variation in unobserved amenities (Sorkin, 2018), making firm fixed effects a reasonable proxy for amenities. I then drop the top two and bottom two ventiles, since the variance of the firm fixed effect is much higher as we move into the tails of the distribution. Unobservable amenities can be reasonably assumed to be roughly constant within the remaining ventiles. I then re-estimate the firm specification defining the cell as a ventile of the distribution of wage fixed effects and report the results in Panel B of Table A.3. The evidence is similar to what was observed when grouping firms by sector in Section 5.1. Tipping points are present in some wage fixed effect ventile cells; the evidence is again strongest for the earlier part of the sample,

[20] The wage effects are estimated on the full sample of workers and firms subject to social security and included directly as a variable in the BHP from 1985, see Bellmann et al. (2020) for details of the estimation.

when net migration was positive, and in particular the period 1990–1995, where tipping points are identified for six out of 16 cells.

I also consider how the evidence for tipping points is affected by using finer geographic groupings, specifically, local labour markets. Again, this does not change the evidence in favour of tipping points, which are identified in 15 per cent of local labour markets over time. Interestingly, however, tipping points are more clearly present during periods of positive net migration, particularly 1980–1995, than when using coarser geographic groupings, where identified tipping points are more evenly spread over time.[21]

5.2.2 Alternate units of analysis

Production teams: I also consider whether the firm is the right level of analysis. One might contend that the correct level of analysis is in fact the production team, not the firm, since it is within such teams that the interpersonal interactions with immigrants in which natives may experience disutility take place. Tipping in the composition of production teams might lead to sorting across production teams within the firm, without necessarily leading to observable tipping dynamics in the overall composition of the firm.

Since I do not observe information on individual workers' occupations or on the composition of firms, I cannot directly test for tipping points at the sub-firm level. However, to establish that the firm is not too large a unit of analysis, I repeat my main estimation specification, limiting the sample to small and medium-sized firms, i.e. those firms employing 10–49 workers. I report the results of these specifications in Table A.4. The evidence in favour of the existence of tipping points is very similar to when considering all firms, particularly when grouping firms by sector or regional labour market, both on average and in different years. For example, I continue to identify tipping points in 40 per cent of sector cells in 1990–1995, as was the case when using all firms. When grouping firms by the intersection of labour market and skill, there is slightly more evidence of tipping points in small firms, particularly in 1990–1995, when tipping points are identified in 24 per cent of cells. However, the evidence from small firms does not alter the conclusion that tipping points are likely to exist, but only in specific industries and in years when there is a sufficiently large shock to the relative supply of immigrant workers.

Industries: As noted previously, the presence or absence of tipping points in the composition of firms does not necessarily rule in or out the possibility that these might be present in the composition of industries. To test for the presence of tipping points in the composition of industries, I estimate Equation (4) over local industries, i.e. three-digit industries by local labour markets, using the same third-order polynomial specification

[21] Note that it would not be feasible to repeat the threshold regressions on local industries, which are defined as the intersection of three-digit industry and local labour market, when grouping by either three-digit industry or local labour market as was done for firms in this section, as the resulting cells would contain too few observations.

with an intercept shift and including controls for log median native wage in the industry, share of low-skilled employment, average firm size, and the Herfindahl-Hirschman index of employment concentration in the local industry. The regressions are again run separately for each cell. I report the results in Table A.5.

In Panel A I report the results when allowing the location of the tipping point to vary by industrial sector. The average estimated tipping point corresponds to a base-year immigrant share between 10 and 16 per cent, somewhat lower than for firms and reflecting the fact that variance in immigrants shares is on average lower in local industries than in firms. The average NLS estimate of the discontinuity in net native employment growth is positive in most years and a negative and significant discontinuity is identified in only 8 per cent of sector-years, reflecting fairly weak support for the existence of sector-specific tipping points in the composition of local industries. Furthermore, the correlation between whether a tipping point is identified for firms in a sector and for industries in the same sector is moderately positive, at 0.2. There does not appear to be strong evidence of sector-specific tipping points in local industry composition over and above any tipping points that might exist in firm composition.

In Panels B and C I report results when grouping local industries by either regional labour markets or skill-type (high or low skill) by regional labour market. The average estimated discontinuity is more often negative and, across years, it is significantly negative in 14 per cent of labour market-years and 16 per cent of skill by labour market-years. However, the threshold model is estimated using relatively few observations when grouping local industries in this way, typically 85–150, so the results should perhaps be interpreted with caution; in the case of industry-labour market cells there are frequently too few observations to estimate a threshold model for a given cell. Median upper and lower bounds for a 95 per cent confidence interval reflect the decrease in precision of the estimates. Finally, the correlation across either regional labour markets or region-skill types cells where tipping points are observed for firms and where they are observed for local industries is of the order of -0.05. This suggests that the admittedly limited evidence for tipping points in industry composition observed when grouping local industries by location does not simply reflect tipping points in the composition of the underlying firms.

6 Interpretation

6.1 Preference spillovers and amenities

The results presented in Section 5 show that tipping points are most likely to exist in low-skill sectors and are not likely to be a widely-shared characteristic of firms. This relatively limited and localised evidence of tipping points in the composition of firms by nativity contrasts with the results presented by Pan (2015), who finds strong evidence for the existence of tipping points in the composition of occupations, even when grouping occupations by quite coarse region-skill cells. This is in spite of a long-standing body of research documenting differences in amenities across occupations going back to Lucas (1977) and Brown (1980) and clear gender differences in the preference for different amenities (Bell, 2022; Goldin, 2014a; Mas and Pallais, 2017).

One difference between the two settings that might explain the divergent results is that the entry of women into the labour market during the second half to the twentieth century is a much larger shock to the relative supplies of types of workers than the inflows of immigrants studied here. To asses whether evidence of tipping points would be more widespread across the labour market in the event of a sufficiently large inflow of immigrants, I extend my analysis to a more recent case study, since the immigrant population in Germany increased from six million in 2011 to 11 million in 2021. Specifically, I consider the evidence for tipping points during the period 2013–2018 for firms located in regional labour markets that experienced an above-median net immigrant inflow, normalised by the total population in 2013, in either East or West Germany. The results are presented in Table A.6. Tipping points are identified for an above-average share of industry or geographic cells when compared to the results in Section 5.1. This suggests that sufficiently large inflows are indeed needed to start to observe tipping in the composition of some firms. However, the overall picture is not dramatically different from other high-tipping point periods such as 1990–1995; evidence of tipping points remains confined to a relatively small subset of firms, operating in low-skill sectors such as Transport, storage, and communication, or Mining.

The fact that tipping points are confined to less-attractive industries, even in the event of a large immigrant inflow, is consistent with other firm-specific amenities playing a more important role than the immigrant share in shaping worker sorting over firms, just as Banzhaf and Walsh (2013) have shown theoretically in the case of neighbourhood sorting. In the theory presented in Section 2, the failure to detect evidence of a tipping point in an industry is consistent with preference spillovers being relatively weak, such that, for most firms in the industry, $\partial \omega^N / \partial n_N > \partial \omega^N / \partial s$ over the entire range of the immigrant share s. In this case, the wage schedule for natives in Figure 2 would be downward sloping over the entire range; the firm would have a single integrated equilibrium that will adjust continuously in response to shifts in the relative supply of immigrants. It is plausible that this is more likely to be the case for firms with amenities natives particularly value, since positive amenities increase the number of natives willing to work at a firm for a given wage and presumably make them less sensitive to the immigrant share in the firm, which is a type of disamenity.

The finding that tipping points are not endemic, but rather are only present in firms operating in certain industries is similar to recent evidence on the importance of tipping points in the racial composition of schools. Using different methods to the ones used here, Caetano and Maheshri (2023) find that endogenous dynamic responses to the racial composition of schools only play a small part in explaining observed trends in school segregation. Simulations show that, across years and school types, the share of schools that have a tipping point their racial composition at all, regardless of whether tipping is actually observed in reality, ranges from as low as 0.2 to 0.6 (Caetano and Maheshri, 2017).

6.2 The role of preference spillovers in explaining segregation

The limited evidence of tipping points presented here suggests that worker preferences, and the preference spillovers that are a necessary condition for the existence of tipping points, are not the main cause underlying observed cross-sectional patterns of workplace segregation. A recently developing literature on the causes of workplace segregation has tended to focus on the importance of homophily in hiring patterns for understanding segregation, given than immigrant managers are disproportionately likely to hire immigrant workers (Åslund et al., 2014). To give a fuller picture of the relative importance of the different causes of segregation, I now turn to changes in aggregate segregation over time.

If segregation is entirely due to managers hiring conationals, recent research would lead us to expect that segregation is non-increasing over time, at least for a given cohort of firms. Kerr and Kerr (2021) have found that firms in the US founded by immigrants continue to disproportionally hire co-ethnics even as the firm ages, but that the tendency to hire co-ethnics does not increase, while Miller and Schmutte (2021) have found that, while new hires in Brazilian firms are disproportionally of the race of the founder, this difference fades out as hires accumulate, which would lead to decreasing segregation over time. Job search on social networks by immigrants and the use of referrals in hiring (e.g. Dustmann et al., 2016; Miller and Schmutte, 2021), while similarly leading to persistence in the composition of the firm, would also ultimately lead to decreasing segregation, so long as there is still some random component to separations and hiring. In contrast, if there are tipping points in firm composition, segregation might be non-decreasing, or even increasing over time, since segregated equilibria in firm composition will tend to be more persistent than integrated equilibria.[22]

To study the dynamics of segregation, I separate the firms in my dataset into five-year cohorts based on their founding date, starting from 1975, so the first cohort is firms founded in 1975–1980, the second cohort is firms founded in 1980–1985, and so forth. I then re-calculate both the index of coworker segregation and the effective index of coworker segregation (Hellerstein and Neumark, 2008), presented in Section 3, separately for the workers of firms in each cohort.[23] The cohort-year-specific index of coworker segregation is reported in Figure 5.

Segregation in a given cohort decreases over time and segregation in a given year decreases with cohort age, as shown in Figure 5a, which is again inconsistent with widespread tipping dynamics. For example, the probability that immigrant workers at firms in the 1975–1980 cohort have of working with other immigrants, relative to the probability natives at these firms have of working with immigrants is around 16 percentage points higher than would occur under a random assignment of workers to firms in 1985, but only 8 percentage points higher in 2005. However, a large part of this pattern is due to the most segregated firms exiting the labour market, as shown by focusing on firms that survive to 2005, in Figure 5b. When focusing on survivors, younger firms still tend to be more segregated in any given year, however existing firms become more segregated in periods where there is a large inflow of immigrants, in 1990–1995 and, to a lesser extent, 1995–2000. The same pattern can be observed in the index of effective coworker segregation, shown in Figure A.1;[24] these patterns are the result of changes to the distribution of the immigrant share across firms, as shown in Figure A.2, rather than a result of differential growth in more- or less-segregated firms.

The evidence presented here on the life cycle of firms is consistent with the findings of, e.g., Miller and Schmutte (2021) that new hires and, as a result, the firm as a whole, tend to become less segregated over time, at least in normal times. However the clear increase in segregation that occurs when there is a large inflow of immigrants suggests an interplay between preference spillovers and hiring practices. The tipping points documented in

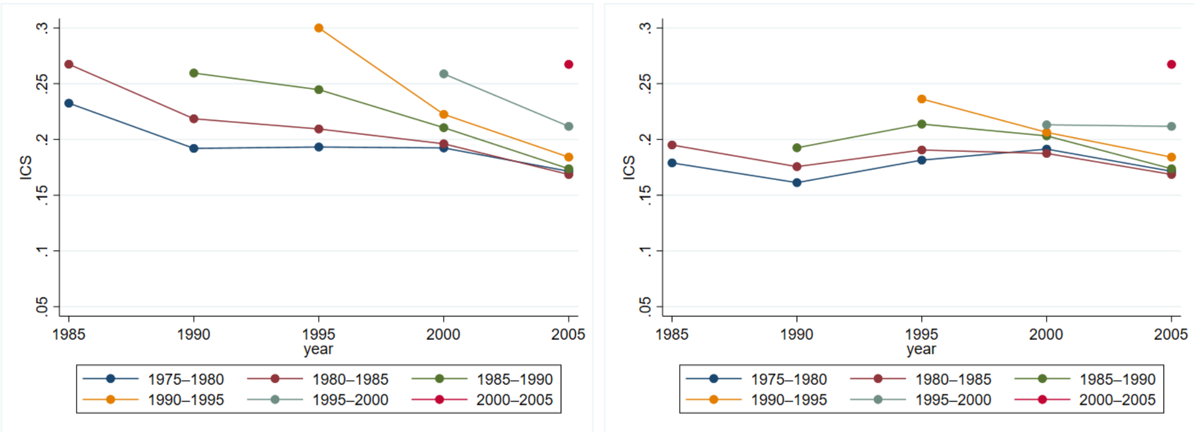
[22]Consider the model presented in Section 2. After an increase in the relative supply of immigrants, causing the firm to tip from an integrated equilibrium, labelled E_2 in Figure 3, future decreases in the relative supply of immigrants will need to be very large for the firm to start hiring natives again. Specifically, it would require that $\omega^I(s, s) > \omega^N(1 - s, s)$ for large values of s, i.e. immigrants become so scarce in the labour market that they need to be paid more to work in a high-migrant firm than a native does, in spite of natives' aversion to working with high shares of immigrants. Similarly, Schelling-style models of tipping model the current minority share as an S-shaped function of the past minority share; this defines a difference equation with one unstable, integrated equilibrium, and two stable, segregated equilibria, see, e.g.,Caetano and Maheshri (2023).

[23]When simulating the counterfactual distribution of workers under random assignment to calculate the effective index, I condition the share of immigrants on labour market and three-digit industry, but not on firm cohort.

[24]The share of immigrants in a labour market by industry in a given year used to calculate the counterfactual random distribution of immigrants across firms are calculated using all firms, not only firms still in operation in 2005.

Figure 5: Index of Coworker Segregation by cohort

(a) All firms (b) Survivors in 2005



Notes: Index of coworker segregation (Hellerstein and Neumark, 2008), calculated separately for workers employed by firms of different cohorts. The number of observations used to calculated each cohort-year value is reported in Tables A.7 and A.8.

Section 5.1, particularly in the period 1990–1995, occur at relatively high immigrant shares. The process of gradual desegregation described by Miller and Schmutte (2021), combined with the natural churn in a firm's workforce, would ensure that the share of immigrants in a firm typically stays comfortably below the tipping point in normal times. However, in the event of a large immigrant inflow, the immigrant share may increase more quickly in some firms, particularly in firms with immigrant managers. Some of these firms may have a tipping point which, if the inflow of newly hired immigrants is large enough, they will cross, reinforcing the increase in segregation caused by the immigrant inflow and homophily in hiring alone. Preference spillovers would, in this case, not be an important cause of observed segregation in normal times, however they would contribute to the increase in segregation that is observed in the event of large immigrant inflows.

7 Conclusion

Tipping-like dynamics have been identified in neighbourhood composition, school enrolments or occupational composition. This paper considered whether tipping points also exist in the composition of workplaces by nativity. Similar to the latest findings in schools (Caetano and Maheshri, 2017, 2023) or neighbourhoods (Caetano and Maheshri, 2021), notwithstanding differences in the methods used to test for tipping dynamics, and distinct from earlier work on occupational segregation in the labour market (Pan, 2015), I find only limited evidence of tipping points in the composition of firms. This evidence is strongest in years where immigrant inflows are largest and for firms operating in low-skill, low-wage sectors. Preference spillovers are therefore likely to make at best a modest contribution to observed patterns of workplace segregation.

Descriptive evidence on the patterns of segregation across firms over time suggests that preference spillovers may be reinforcing increases in segregation caused in years of large immigrant inflows by manager hiring practices and job search behaviour. However, in normal times, segregation tends to decrease over firms, as has been documented in other settings. A productive avenue for future research would be to jointly consider and quantify the role of manager hiring practices and job search on networks, firm-level amenities, and preferences spillovers for determining the distribution of workers of different nativities over firms and for determining where immigrants find work, in particular in periods of large immigrant inflows.

References

Aldén, L., M. Hammarstedt, and E. Neuman (2015): "Ethnic segregation, tipping behavior, and native residential mobility," *International Migration Review*, 49, 36–69.

Andersson, F., M. García-Pérez, J. Haltiwanger, K. McCue, and S. Sanders (2014): "Workplace Concentration of Immigrants," *Demography*, 51, 2281–2306.

Andrews, I., T. Kitagawa, and A. McCloskey (2019): "Inference on Winners," Working Paper 25456, National Bureau of Economic Research, Boston.

——— (2021): "Inference after estimation of breaks," *Journal of Econometrics*, 224, 39–59.

Åslund, O., L. Hensvik, and O. N. Skans (2014): "Seeking Similarity: How Immigrants and Natives Manage in the Labor Market," *Journal of Labor Economics*, 32, 405–441.

Åslund, O. and O. N. Skans (2010): "Will I see you at work: Ethnic Workplace Segregation in Sweden, 1985-2002," *ILR Review*, 63, 471–493.

Aydemir, A. and M. Skuterud (2008): "The immigrant wage differential within and across establishments," *Industrial and Labor Relations Review*, 61, 334–352.

Banzhaf, H. S. and R. P. Walsh (2013): "Segregation and Tiebout sorting: The link between place-based investments and neighborhood tipping," *Journal of Urban Economics*, 74, 83–98.

Barth, E., B. Bratsberg, and O. Raaum (2012): "Immigrant wage profiles within and between establishments," *Labour Economics*, 19, 541–556.

Becker, G. S. and K. M. Murphy (2000): *Social Economics: Market Behaviour in a Social Environment*, Harvard University Press.

Bell, A. (2022): "Job Amenities and Earnings Inequality," Available at SSRN: https://ssrn.com/abstract=4173522.

Bellmann, L., B. Lochner, S. Seth, and S. Wolter (2020): "AKM effects for German labour market data," FDZ Methodenreport 01|2020, Institute for Employment Research (IAB), Nürnberg.

Böhlmark, A. and A. Willén (2020): "Tipping and the effects of segregation," *American Economic Journal: Applied Economics*, 12, 318–347.

Brinatti, A. and N. Morales (2021): "Firm Heterogeneity and the Impact of Immigration: Evidence from German Establishments," Working Paper 21-16, Federal Reserve Bank of Richmond.

Brown, C. (1980): "Equalizing differences in the labor market," *Quarterly Journal of Economics*, 94, 113–134.

Caetano, G. and V. Maheshri (2017): "School segregation and the identification of tipping behavior," *Journal of Public Economics*, 148, 115–135.

——— (2021): "A Unified Empirical Framework to Study Segregation," Available from https://sites.google.com/site/vmaheshri/papers.

——— (2023): "Explaining Recent Trends in US School Segregation," *Journal of Labor Economics*, Forthcoming.

Card, D., A. Mas, and J. Rothstein (2008): "Tipping and the dynamics of segregation," *Quarterly Journal of Economics*, 123, 177–218.

——— (2011): "Are mixed neighborhoods always unstable? two-sided and one-sided tipping," in *Neighborhood and Life Chances: How Place Matters in Modern America*, ed. by H. B. Newburger, E. L. Birch, and S. M. Wachter, University of Pennsylvania Press, chap. 14.

Dustmann, C., A. Glitz, U. Schönberg, and H. Brücker (2016): "Referral-based job search networks," *Review of Economic Studies*, 83, 514–546.

Elliott, G. and U. K. Müller (2007): "Confidence sets for the date of a single break in linear time series regressions," *Journal of Econometrics*, 141, 1196–1218.

European Commission (2018): "Results of special Eurobarometer on integration of immigrants in the European Union," Special Eurobarometer 469.

Ganzer, A., L. Schmidtlein, J. Stegmaier, and S. Wolter (2021): "The Establishment History Panel 1975–2019 (v2)," FDZ Datenreport 16|2020, Institute for Employment Research (IAB), Nürnberg.

Glitz, A. (2014): "Ethnic segregation in Germany," *Labour Economics*, 29, 28–40.

Glover, D., A. Pallais, and W. Pariente (2017): "Discrimination as a self-fulfilling prophecy: evidence from French grocery stores," *Quarterly Journal of Economics*, 132, 1219–1260.

Goldin, C. (2014a): "A grand gender convergence: Its last chapter," *American Economic Review*, 104, 1091–1119.

——— (2014b): "A pollution theory of discrimination: Male and female differences in occupations and earnings," in *Human Capital in History: The American Record*, ed. by L. P. Boustan, C. Frydman, and R. A. Margo, University of Chicago Press, 313–348.

Hansen, B. E. (1996): "Inference when a nuisance parameter is not identified under the null hypothesis," *Econometrica*, 64, 413–430.

——— (2011): "Threshold autoregression in economics," *Statistics and its Interface*, 4, 123–128.

——— (2021): *Econometrics*, Princeton University Press.

Hellerstein, J. K. and D. Neumark (2008): "Workplace Segregation in the United States: Race, ethnicity, and skill," *The Review of Economics and Statistics*, 90, 459–477.

Higgs, R. (1977): "Firm-Specific Evidence on Racial Wage Differentials and Workforce Segregation," *American Economic Review*, 67, 236–245.

Hjort, J. (2014): "Ethnic divisions and production in firms," *Quarterly Journal of Economics*, 129, 1899–1946.

Hsieh, C.-T., E. Hurst, C. I. Jones, and P. J. Klenow (2019): "The allocation of talent and U.S. economic growth," *Econometrica*, 87, 1439–1474.

Imbens, G. W. and T. Lemieux (2008): "Regression discontinuity designs: A guide to practice," *Journal of Econometrics*, 142, 615–635.

Kerr, S. P. and W. R. Kerr (2021): "Whose Job Is It Anyway? Co-Ethnic Hiring in New U.S. Ventures," NBER Working Paper 28509.

Kropp, P. and B. Schwengler (2011): "Abgrenzung von Arbeitsmarktregionen – ein Methodenvorschlag," *Raumforschung und Raumordnung*, 69, 45–62.

Lee, D. S. and T. Lemieux (2010): "Regression discontinuity designs in economics," *Journal of Economic Literature*, 48, 281–355.

Lepage, P.-L. (2021): "Endogenous Learnings, Persistent Employer Biases, and Discrimination," Available at SSRN: https://ssrn.com/abstract=3640663.

Lubotsky, D. H. (2007): "Chutes or Ladders? A Longitudinal Analysis of Immigrant Earnings," *Journal of Political Economy*, 115, 820–867.

Lucas, R. E. B. (1977): "Hedonic wage equations and psychic wages in the returns to schooling," *American Economic Review*, 67, 549–558.

Mas, A. and A. Pallais (2017): "Valuing alternative work arrangements," *American Economic Review*, 107, 3722–3759.

Miller, C. and I. Schmutte (2021): "The Dynamics of Referral Hiring and Racial Inequality: Evidence from Brazil," NBER Working Paper 29246.

Noiriel, G. (1988): *Le creuset français: Histoire de l'immigration (XIXe-XXe siècle)*, Paris: Éditions du Seuil.

Oberfield, E. and D. Raval (2021): "Micra data and macro technology," *Econometrica*, 89, 703–732.

OECD (2020): "International migration database," Data retrieved from OECD International Migration Statistics, https://doi.org/10.1787/mig-data-en.

Ottaviano, G. I. and G. Peri (2012): "Rethinking the effect of immigration on wages," *Journal of the European Economic Association*, 10, 152–197.

Pan, J. (2015): "Gender segregation in occupations: The role of tipping and social interactions," *Journal of Labor Economics*, 33, 365–408.

Rho, D. and S. Sanders (2021): "Immigrant earnings assimilation in the united states: A panel analysis," *Journal of Labor Economics*, 39, 37–78.

Schelling, T. C. (1971): "Dynamic models of segregation," *Journal of Mathematical Sociology*, 1, 143–186.

——— (1978): *Micromotives and Macrobehavior*, New York: Norton.

Sorkin, I. (2018): "Ranking firms using revealed preference," *Quarterly Journal of Economics*, 133, 1331–1393.

Zheng, L. (2014): "Essays in Labor Economics," Ph.D. thesis, Cornell University.

A Supplementary tables and figures

Table A.1: Sectors where tipping points are identified

	1975 (1)	1980 (2)	1985 (3)	1990 (4)	1995 (5)	2000 (6)	2005 (7)
A - Agriculture, hunting and forestry	0	0	0	1	0	0	0
C - Mining and quarying	0	0	0	1	0	0	0
D - Manufacturing	1	0	1	0	0	0	0
E - Electricity, gas and water supply	0	0	0	0	0	0	0
F - Construction	0	0	0	0	0	0	0
G - Wholesale and retail trade; repairs	1	0	0	0	0	0	0
H - Hotels and restaurants	0	1	0	1	0	0	1
I - Transport, storage and communication	0	1	0	1	0	0	0
J - Financial intermediation	0	0	0	0	1	0	0
K - Real estate, renting and business activities	0	1	0	0	0	0	0
L - Public administration and defence; social security	0	0	0	0	0	0	0
M - Education	0	0	0	0	0	0	0
N - Health and social work	0	0	1	0	0	0	0
O - Other community, social and personal service activities	0	0	0	1	0	0	1
P - Private households with employed persons	0	1	0	1	.	0	0

Note: Single-letter sectors (NACE Rev. 1) where a negative and significant discontinuity is identified in normalised native workforce growth according to the threshold regression specification in Equation (4).

Table A.2: Labour markets where tipping points are identified

	1975 (1)	1980 (2)	1985 (3)	1990 (4)	1995 (5)	2000 (6)	2005 (7)
Hamburg	0	1	0	0	1	0	1
Braunschweig/Wolfsburg	0	0	0	0	0	0	0
Göttingen	0	0	0	0	0	1	0
Hannover	0	0	0	0	0	0	0
Oldenburg(O.)	0	0	0	0	0	0	1
Osnabrück	0	0	0	0	0	0	0
Bremen	1	0	0	0	0	0	0
Düsseldorf-Ruhr	0	0	0	0	0	0	1
Aachen	1	1	0	0	0	0	0
Köln	0	0	0	1	1	0	0
Münster	0	0	0	1	1	0	0
Bielefeld/Paderborn	0	1	0	0	0	1	0
Siegen	0	0	1	0	0	1	1
Frankfurt a.M.	0	0	0	1	0	0	0
Kassel	0	0	0	0	0	0	0
Koblenz	0	1	0	0	0	0	0
Trier	0	0	0	0	0	0	0
Stuttgart	0	0	0	0	0	0	1
Karlsruhe	0	0	0	0	0	0	1
Mannheim	0	0	0	0	0	0	0
Freiburg i.Br.	0	0	0	0	1	0	0
Offenburg	0	0	0	0	1	0	0
Villingen-Schwenningen	0	0	0	0	0	1	0
Konstanz	0	0	0	0	0	0	0
Lörrach	0	0	0	1	0	0	0
Ulm	1	0	1	0	1	0	0
Ravensburg	0	0	0	0	0	0	1
München	0	0	0	1	1	1	0
Passau	0	0	0	0	0	0	0
Regensburg	0	0	1	0	0	0	0
Weiden i.d.OPf.	0	0	0	0	0	0	0
Bayreuth	0	1	0	0	0	0	0
Coburg	0	1	1	0	0	0	0
Hof	0	0	0	0	0	0	0
Wunsiedel i.F.	0	0	0	0	1	0	0
Nürnberg	0	0	0	0	0	0	0
Schweinfurt	0	0	0	1	0	0	0
Würzburg	0	0	1	0	0	0	0
Saarbrücken	0	0	0	1	0	0	0

Note: Regional labour markets (Kropp and Schwengler, 2011) where a negative and significant discontinuity is identified in normalised native workforce growth.

Table A.3: Alternative groupings of firms

	(1) 1975	(2) 1980	(3) 1985	(4) 1990	(5) 1995	(6) 2000	(7) 2005
A: 3-digit industry							
Tipping point	23.7	23.4	19.1	21.2	23.0	20.6	23.6
	(19.7)	(19.7)	(17.5)	(20.7)	(21.6)	(19.7)	(21.2)
Discontinuity ($\hat{\delta}$)	1139.8	-4.4	7.7	-1.8	4.8	6.0	66.9
	(13356.3)	(68.2)	(278.7)	(71.5)	(108.3)	(71.6)	(774.0)
$\hat{\delta} < 0$ and p-val. < 0.05	0.17	0.21	0.12	0.15	0.16	0.18	0.11
Median LB, 95% CI	-51.74	-46.02	-46.01	-46.91	-44.98	-46.29	-28.45
Median UB, 95% CI	46.01	40.56	42.24	39.32	55.08	48.21	59.45
Cells	136	141	145	149	151	157	151
Median obs.	247	240	228	252	201	328	387
B: wage FE ventile							
Tipping point			51.5	45.7	56.1	45.5	33.3
			(33.0)	(28.0)	(28.3)	(36.8)	(22.6)
Discontinuity ($\hat{\delta}$)			16.1	-8.8	-29.1	-4.2	-1.1
			(84.9)	(66.9)	(78.2)	(62.0)	(25.2)
$\hat{\delta} < 0$ and p-val. < 0.05			0.13	0.38	0.25	0.19	0.06
Median LB, 95% CI			-8.35	-11.55	-61.72	-21.30	-24.09
Median UB, 95% CI			25.68	30.78	29.20	15.08	20.99
Cells			16	16	16	16	16
Median obs.			6025	6666	5171	7280	7782
C: Local labour market							
Tipping point	24.5	26.1	21.4	22.9	28.2	27.0	28.0
	(20.1)	(22.8)	(18.2)	(20.7)	(23.2)	(22.7)	(22.4)
Discontinuity ($\hat{\delta}$)	13.8	1.2	-11.9	1.8	16.2	11.9	20.0
	(73.5)	(80.4)	(81.7)	(62.2)	(93.6)	(68.5)	(67.4)
$\hat{\delta} < 0$ and p-val. < 0.05	0.12	0.21	0.21	0.16	0.15	0.13	0.07
Median LB, 95% CI	-35.46	-38.30	-49.98	-31.90	-39.09	-32.29	-28.24
Median UB, 95% CI	46.44	30.50	35.57	35.95	50.06	33.85	48.47
Cells	85	86	86	86	86	86	86
Median obs.	617	669	657	734	766	1026	1085

Note: Summary of a set of threshold regressions. In panel A firms are grouped by three-digit industry (NACE Rev. 1), in Panel B firms are grouped by fixed effect ventiles from a worker-firm wage regression (Bellmann et al., 2020), available from 1985 on, in Panel C firms are grouped by local labour market (Kropp and Schwengler, 2011). Inference is conducted using the methods proposed by Andrews et al. (2021).

Table A.4: Tipping points in the composition of small firms

	(1) 1975	(2) 1980	(3) 1985	(4) 1990	(5) 1995	(6) 2000	(7) 2005
A: Industrial sector							
Tipping point	31.9	33.1	34.9	50.4	44.5	34.8	37.4
	(27.9)	(31.8)	(36.4)	(37.6)	(34.3)	(31.3)	(33.1)
Discontinuity ($\hat{\delta}$)	11.8	-27.0	-8.2	-26.2	-37.1	7.7	9.2
	(59.7)	(57.8)	(62.2)	(82.0)	(96.4)	(54.8)	(46.8)
$\hat{\delta} < 0$ and p-val. < 0.05	0.20	0.20	0.13	0.40	0.14	0	0
Median LB, 95% CI	-22.12	-38.97	-39.20	-60.64	-29.32	-0.62	-14.51
Median UB, 95% CI	31.41	15.78	20.70	19.66	30.52	58.32	47.32
Cells	15	15	15	15	14	15	15
Median obs.	2751	3273	3692	3926	3934	6733	7163
B: Regional labour market							
Tipping point	26.9	33.3	23.8	28.3	30.5	33.3	29.7
	(19.5)	(25.4)	(24.0)	(25.5)	(25.1)	(28.8)	(27.3)
Discontinuity ($\hat{\delta}$)	12.6	13.3	7.3	-10.6	12.0	-3.0	4.3
	(74.8)	(79.5)	(65.4)	(69.6)	(88.0)	(81.3)	(57.5)
$\hat{\delta} < 0$ and p-val. < 0.05	0.10	0.13	0.18	0.18	0.21	0.05	0.21
Median LB, 95% CI	-45.91	-25.37	-37.38	-42.48	-34.76	-33.16	-25.27
Median UB, 95% CI	68.11	41.30	36.95	29.03	50.76	45.50	19.89
Cells	39	39	39	39	39	39	39
Median obs.	865	989	1034	1142	1206	1540	1652
C: Region-sector type							
Tipping point	24.3	28.7	20.3	22.6	27.4	30.5	29.4
	(19.4)	(25.4)	(18.4)	(19.6)	(23.2)	(24.4)	(24.4)
Discontinuity ($\hat{\delta}$)	-64.3	-42.3	40.8	-17.1	20.5	11.0	-5.0
	(536.9)	(373.0)	(213.8)	(137.2)	(94.8)	(94.1)	(65.4)
$\hat{\delta} < 0$ and p-val. < 0.05	0.14	0.16	0.08	0.24	0.12	0.14	0.18
Median LB, 95% CI	-53.81	-54.48	-31.17	-47.42	-34.22	-39.71	-39.25
Median UB, 95% CI	54.39	45.97	62.05	38.33	62.12	49.44	42.07
Cells	72	73	74	78	78	78	78
Median obs.	471	537	515	519	547	714	776

Note: Summary statistics for a set of threshold regressions. In all cases the sample has been restricted to firms employing 10-49 workers in the base year. In Panel A each regression uses firms from a given single-letter industrial sector (NACE Rev. 1), in Panel B a regression uses firms from a regional labour market Kropp and Schwengler (2011), in Panel C a regression uses firms of a given skill level (high or low) in a given labour market. Inference is conducted using the methods proposed by Andrews et al. (2021).

Table A.5: Tipping points in the composition of local industries

	(1) 1975	(2) 1980	(3) 1985	(4) 1990	(5) 1995	(6) 2000	(7) 2005
A: Industrial sector							
Tipping point	11.1	10.8	9.6	8.4	8.7	13.4	9.4
	(6.8)	(7.4)	(3.6)	(4.0)	(5.0)	(11.4)	(5.1)
Discontinuity ($\hat{\delta}$)	6.3	5.1	6.3	8.8	2.8	21.1	-1.9
	(46.4)	(22.0)	(22.1)	(28.8)	(31.5)	(52.4)	(31.1)
$\hat{\delta} < 0$ and p-val. < 0.05	0.15	0	0.08	0	0.08	0.15	0.17
Median LB, 95% CI	-3.96	-11.62	-29.26	-14.33	-41.34	-9.80	-21.97
Median UB, 95% CI	25.97	25.12	28.70	26.61	42.46	36.57	13.19
Cells	13	12	12	13	13	13	12
Median obs.	263	280	287	277	285	357	381
B: Regional labour market							
Tipping point	9.8	11.0	9.4	10.5	12.7	11.6	10.4
	(5.3)	(6.4)	(3.9)	(5.2)	(7.2)	(9.2)	(5.4)
Discontinuity ($\hat{\delta}$)	58.4	-10.1	-126.5	-1.0	38.5	15.5	13.5
	(255.0)	(56.8)	(748.1)	(46.6)	(106.8)	(103.4)	(61.6)
$\hat{\delta} < 0$ and p-val. < 0.05	0.13	0.25	0.13	0.16	0.08	0.22	0.03
Median LB, 95% CI	-34.02	-50.32	-38.82	-49.43	-31.78	-36.21	-19.78
Median UB, 95% CI	45.68	31.75	45.36	34.16	116.15	50.67	44.85
Cells	32	32	30	32	37	37	35
Median obs.	104	113	115	126	113	141	146
C: Region-sector type							
Tipping point	10.5	10.6	9.8	10.5	12.7	10.6	11.4
	(6.7)	(8.7)	(4.4)	(6.7)	(7.7)	(6.2)	(5.4)
Discontinuity ($\hat{\delta}$)	13.4	6.7	-111.3	4.7	5.1	-333.8	9.5
	(64.0)	(69.4)	(726.2)	(53.0)	(92.9)	(2434.9)	(68.5)
$\hat{\delta} < 0$ and p-val. < 0.05	0.14	0.18	0.22	0.16	0.16	0.11	0.16
Median LB, 95% CI	-35.81	-33.75	-50.01	-42.00	-62.73	-37.25	-39.64
Median UB, 95% CI	41.56	37.32	43.42	46.34	69.11	50.01	43.22
Cells	36	39	37	43	50	53	55
Median obs.	85	86	89	88	81	91	88

Note: Summary statistics on for a set of threshold regressions where an observation is a local labour market by 3-digit industry. In Panel A each regression uses local industries from a given single-letter industrial sector (NACE Rev. 1), in Panel B a regression uses local industries from a regional labour market Kropp and Schwengler (2011), in Panel C a regression uses local industries of a given skill level (high or low) in a given labour market. Inference is conducted using the methods proposed by Andrews et al. (2021).

Table A.6: Tipping points in 2013–2018

	(1) Industry	(2) Sector	(3) Lab. market	(4) Skill-LM
Tipping point	27.6	49.5	51.4	46.5
	(25.0)	(31.6)	(31.7)	(28.4)
Discontinuity ($\hat{\delta}$)	-3.3	-23.3	14.7	8.5
	(88.0)	(54.0)	(84.3)	(96.3)
$\hat{\delta} < 0$ and p-val. < 0.05	0.18	0.13	0.12	0.24
Median LB, 95% CI	-48.86	-61.04	-24.07	-35.15
Median UB, 95% CI	39.25	12.99	52.09	46.33
Cells	173	15	25	50
Median obs.	254	8062	2775	1434

Note: Summary statistics for a set of threshold regressions. The definition of a cell varies by column. Inference is conducted using the methods proposed by Andrews et al. (2021).

Table A.7: Number of observations for ICS calculations

	1985	1990	1995	2000	2005
1975–1980	87733	76326	63538	60204	47080
1980–1985	92574	80117	64990	62588	50084
1985–1990	0	111930	88578	84483	66838
1990–1995	0	0	123858	118335	91617
1995–2000	0	0	0	218448	160104
2000–2005	0	0	0	0	191155
Total	180307	268373	340964	544058	606878

Note: Tabulates number of establishments employing at least two workers in each cohort-year of observation. Establishments are not required to be observed in all years prior to the observation year, so the apparent size of a cohort can grow over time. Source: BHP.

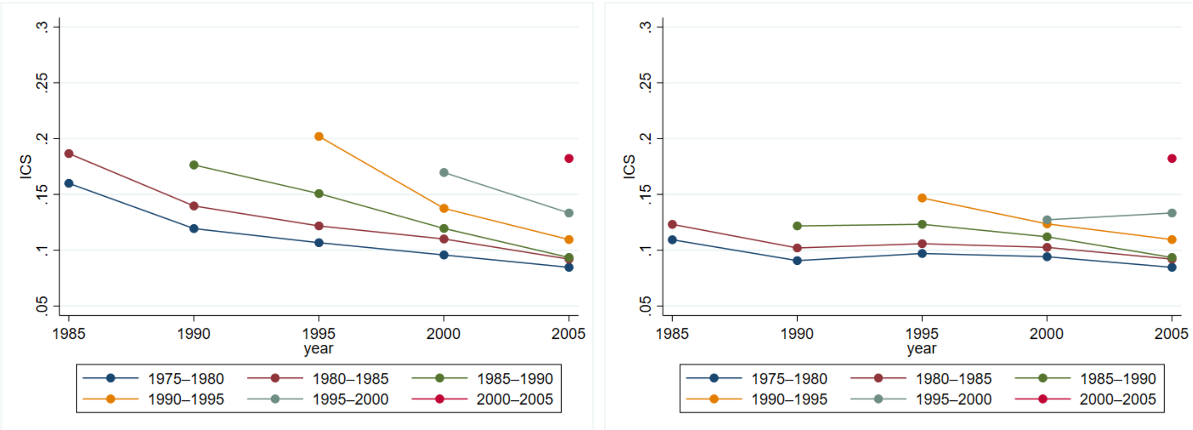
Table A.8: Number of observations for ICS calculations, survivors

	1985	1990	1995	2000	2005
1975–1980	34944	37539	38071	43792	47080
1980–1985	31296	38290	39429	46223	50084
1985–1990	0	43880	50533	61180	66838
1990–1995	0	0	59615	82072	91617
1995–2000	0	0	0	124854	160104
2000–2005	0	0	0	0	191155
Total	66240	119709	187648	358121	606878

Note: Tabulates number of establishments that employ at least two workers and are that are observed in operation in 2005 in each cohort-year of observation. Establishments are not required to be observed in all years prior to 2005, so the apparent size of a cohort can fluctuate over time. Source: BHP.

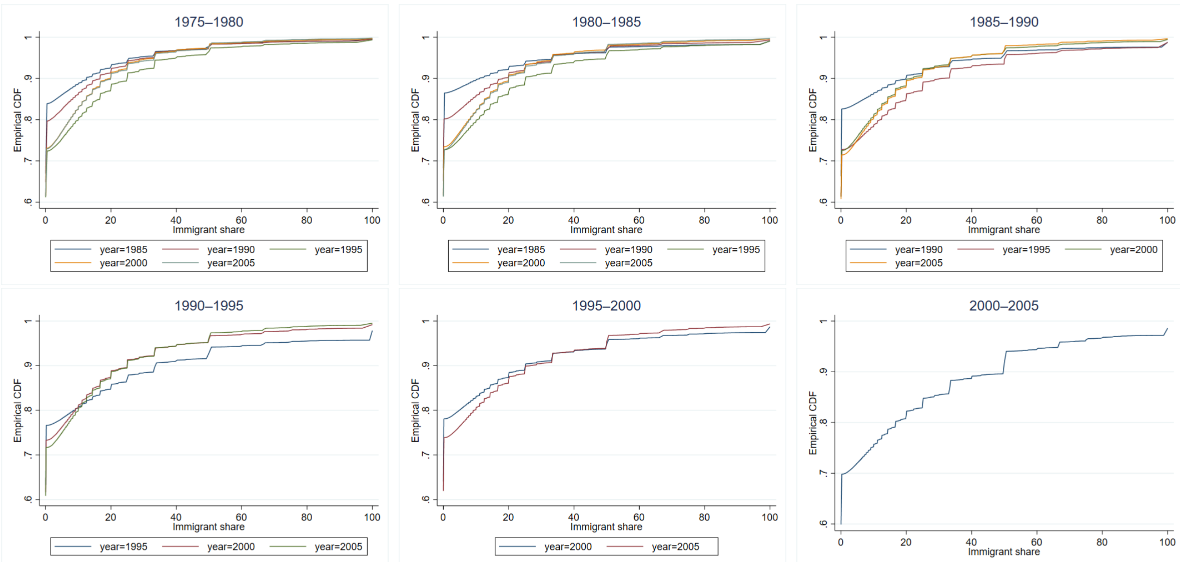
Figure A.1: Index of Coworker Segregation by cohort

(a) All firms (b) Survivors in 2005



Notes: Effective index of coworker segregation (Hellerstein and Neumark, 2008), calculated separately for workers of firms belonging to different cohorts. The counterfactual index is calculated by randomly allocating workers to firms, conditional on the labour market and 3-digit industry of their actual firm. See Tables A.7 and A.8 for the number of observations underlying each estimated index.

Figure A.2: Empirical CDF of immigrant share, firms in operation in 2005



Notes: The figures show the empirical cumulative density function (CDF) of the immigrant share for firms observed in 2005. Each subfigure corresponds to a different cohort.

B Implementation details of inference procedure

This appendix sets out the detail of the threshold model that I estimate and defines the quantities necessary for the implementation of the inference procedures used, which are those developed in Andrews et al. (2019, 2021). The general model I estimate can be written as

$$Y_i = C_i'\beta + D_i'\delta\mathbf{1}(Q_i > \theta_0) + u_i, \tag{B.1}$$

where $C_i \in \mathbb{R}^d$ and $D_i \in \mathbb{R}^l$, with $1 \leq l \leq d$. This is very similar to the set-up considered by Andrews et al. (2021), only I allow for the possibility that the effect of only a sub-vector, D_i, of the full vector of control variables, C_i, varies when the variable Q_i crosses the threshold θ_0. While the results developed by Andrews et al. (2019, 2021) extend straightforwardly to this case, the definitions of various relevant quantities are slightly modified. Here I define the elements necessary to construct the estimators and confidence intervals defined by Andrews et al. (2019, 2021) when estimating the model defined in Equation (B.1).

Consider a finite parameter space Θ. Throughout I will define $\hat{\theta}_n$ as the NLS estimate of θ_0. For all $\theta \in \Theta$ define

$$X_n(\theta) = \begin{pmatrix} \left(\sum_{i=1}^n D_iD_i'\mathbf{1}\{Q_i \leq \theta\}\right)^{-1/2}\left(\sum_{i=1}^n D_i\eta_i\mathbf{1}\{Q_i \leq \theta\}\right) \\ \left(\sum_{i=1}^n D_iD_i'\mathbf{1}\{Q_i > \theta\}\right)^{-1/2}\left(\sum_{i=1}^n D_i\eta_i\mathbf{1}\{Q_i > \theta\}\right) \end{pmatrix} \tag{B.2}$$

where $\eta_i = D_i'\delta\mathbf{1}(Q_i > \theta_0) + u_i$. I assume that the threshold effect, δ, is small relative to sampling variability, which Elliott and Müller (2007) propose to model by assuming that $\delta = n^{-1/2}d$ for some $d \in \mathbb{R}$. Under this assumption, the arguments used in the proof of Proposition (1) in Elliott and Müller (2007) can be applied to show that $\hat{\theta}_n = \text{argmax}_{\theta\in\Theta}\|X_n(\theta)\| + o_p(1)$. This alternative (asymptotic) characterisation of $\hat{\theta}$ is useful to derive asymptotic confidence intervals for $\hat{\theta}_n$ or $\hat{\delta}(\hat{\theta}_n)$. Note furthermore that under the small threshold assumption and standard regularity conditions on the variable moments and covariances, it is straightforward to show that

$$\begin{aligned} X_n(\theta) \xrightarrow{d} X(\theta) = &\begin{pmatrix} \Sigma_{DD}(\theta)^{-1/2}\Sigma_{DDd}(\theta) \\ (\Sigma_{DD}(\bar{\theta}) - \Sigma_{DD}(\theta))^{-1/2})(\Sigma_{DDd}(\bar{\theta}) - \Sigma_{DDd}(\theta)) \end{pmatrix} \\ &+ \begin{pmatrix} \Sigma_{DD}(\theta)^{-1/2}G_D(\theta) \\ (\Sigma_{DD}(\bar{\theta}) - \Sigma_{DD}(\theta))^{-1/2})(G_D(\bar{\theta}) - G_D(\theta)) \end{pmatrix} \end{aligned}$$

where $\bar{\theta} = \sup(\Theta)$ and

$$\begin{aligned} n^{-1}\Sigma_{i=1}^{n} D_i D_i' \mathbf{1}\{Q_i \le \theta\} &\xrightarrow{p} \Sigma_{DD}(\theta) \\ n^{-1}\Sigma_{i=1}^{n} D_i D_i' d\mathbf{1}\{Q_i > \theta_0\}\mathbf{1}\{Q_i \le \theta\} &\xrightarrow{p} \Sigma_{DDd}(\theta) \\ n^{-1/2}\Sigma_{i=1}^{n} D_i u_i \mathbf{1}\{Q_i \le \theta\} &\xrightarrow{d} G_D(\theta) \sim \mathcal{N}(0, \Sigma_{GD}) \end{aligned}$$

Furthermore, define $Y_n(\theta) = e_j\sqrt{n}\hat{\delta}(\theta)$, where $\hat{\delta}(\theta)$ is the OLS estimate of δ after setting $\theta_0 = \theta$ and $e_j \in \mathbb{R}^l$ is the jth basis vector. Then, under the same standard regularity conditions as before, standard regression algebra can be used to show that

$$Y_n(\theta) \xrightarrow{d} \mathcal{A}(\theta)^{-1}(\mathcal{B}(\theta) + \mathcal{C}(\theta)) \tag{B.3}$$

where, extending the previous notation,

$$\begin{aligned} \mathcal{A}(\theta) &= \Sigma_{DD}(\bar{\theta}) - \Sigma_{DD}(\theta) - (\Sigma_{DC}(\bar{\theta}) - \Sigma_{DC}(\theta))\Sigma_{CC}(\bar{\theta})^{-1}(\Sigma_{DC}(\bar{\theta}) - \Sigma_{DC}(\theta))' \\ \mathcal{B}(\theta) &= \Sigma_{DDd}(\bar{\theta}) - \Sigma_{DDd}(\theta) - (\Sigma_{DC}(\bar{\theta}) - \Sigma_{DC}(\theta))\Sigma_{CC}(\bar{\theta})^{-1}\Sigma_{CDd}(\bar{\theta})' \\ \mathcal{C}(\theta) &= G_D(\bar{\theta}) - G_D(\theta) - (\Sigma_{DC}(\bar{\theta}) - \Sigma_{DC}(\theta))\Sigma_{CC}(\bar{\theta})^{-1}G_C(\bar{\theta}). \end{aligned}$$

$X_n(\theta)$ and $Y_n(\theta)$ are therefore asymptotically normal. The asymptotic covariance matrices, $\Sigma_{XY}(\theta, \tilde{\theta})$ and $\Sigma_Y(\theta, \tilde{\theta})$ can be shown to be as follows:

$$\Sigma_{XY}(\theta, \tilde{\theta}) = \begin{pmatrix} \Sigma_{DD}(\theta)^{-1/2}\mathrm{E}[G_D(\theta)\mathcal{C}(\tilde{\theta})']\mathcal{A}(\tilde{\theta})^{-1}e_j \\ (\Sigma_{DD}(\bar{\theta}) - \Sigma_{DD}(\theta))^{-1/2}(\mathrm{E}[G_D(\bar{\theta})\mathcal{C}(\tilde{\theta})'] - \mathrm{E}[G_D(\theta)\mathcal{C}(\tilde{\theta})'])\mathcal{A}(\tilde{\theta})^{-1}e_j \end{pmatrix} \tag{B.4}$$

$$\Sigma_{YY}(\theta, \tilde{\theta}) = e_j'\mathcal{A}(\theta)^{-1}\mathrm{E}[\mathcal{C}(\theta)\mathcal{C}(\tilde{\theta})']\mathcal{A}(\tilde{\theta})^{-1}e_j \tag{B.5}$$

where

$$\begin{aligned} \mathrm{E}[G_D(\theta)\mathcal{C}(\tilde{\theta})'] =& \mathrm{E}[G_D(\theta)G_D(\hat{\theta})'] - \mathrm{E}[G_D(\theta)G_D(\tilde{\theta})'] \\ &- \mathrm{E}[G_D(\theta)G_C(\hat{\theta})']\Sigma_{CC}^{-1}(\Sigma_{DC}(\hat{\theta}) - \Sigma_{DC}(\tilde{\theta}))' \\ \mathrm{E}[\mathcal{C}(\theta)\mathcal{C}(\tilde{\theta})'] =& \mathrm{E}[G_D(\bar{\theta})G_D(\bar{\theta})'] - \mathrm{E}[G_D(\theta)G_D(\bar{\theta})'] \\ &+ (\mathrm{E}[G_D(\theta)G_C(\bar{\theta})'] - \mathrm{E}[G_D(\bar{\theta})G_C(\bar{\theta})'])\Sigma_{CC}(\bar{\theta})^{-1}(\Sigma_{DC}(\bar{\theta}) - \Sigma_{DC}(\tilde{\theta}))' \\ &+ (\Sigma_{DC}(\bar{\theta}) - \Sigma_{DC}(\theta))\Sigma_{CC}(\bar{\theta})^{-1}(\mathrm{E}[G_C(\bar{\theta})G_D(\tilde{\theta})'] - \mathrm{E}[G_C(\bar{\theta})G_D(\bar{\theta})']) \\ &+ (\Sigma_{DC}(\bar{\theta}) - \Sigma_{DC}(\theta))\Sigma_{CC}(\bar{\theta})^{-1}\mathrm{E}[G_C(\bar{\theta})G_C(\bar{\theta})']\Sigma_{CC}(\bar{\theta})^{-1} \\ &\quad \times (\Sigma_{DC}(\bar{\theta}) - \Sigma_{DC}(\tilde{\theta}))'. \end{aligned}$$

The conditional, unconditional, and hybrid confidence intervals and median-unbiased estimators defined in Andrews et al. (2019, 2021) can now be calculated for the model defined in Equation (B.1) by using the definitions of $X(\theta)$, $Y(\theta)$, $\Sigma_{YY}(\theta, \tilde{\theta})$, and $\Sigma_{XY}(\theta, \tilde{\theta})$ derived in this appendix in the definitions of the estimators and confidence intervals proposed by Andrews et al..

When implementing the estimators and confidence intervals defined by Andrews et al. (2021), we replace $X(\theta)$ with $\hat{X}_n(\theta)$, defined in Equation (B.2), where we substitute $\hat{\eta}_i = D_i'\hat{\delta}\mathbf{1}(Q_i > \hat{\theta}_n) + \hat{u}_i$ for η_i, letting, $\hat{\cdot}$ denote the NLS sample estimate of the parameters and errors defined in Equation (B.1). An estimate of $Y(\theta)$ is formed by taking the sample analogue of the limiting random variable in Equation (B.3), i.e. replacing the asymptotic matrices in the definitions of $\mathcal{A}(\theta)$, $\mathcal{B}(\theta)$, and $\mathcal{C}(\theta)$ by their sample analogues. Finally, to estimate the covariance matrices defined in Equations (B.4) and (B.5), I estimate $\mathrm{E}[G_D(\theta)G_D(\tilde{\theta})']$ using the heteroskedasticity-robust sample covariance matrix $n^{-1}\sum_{i=1}^{n} D_i D_i' \hat{u}_i^2 \mathbf{1}\{Q_i \leq \min(\theta, \tilde{\theta})\}$, where $\hat{u}_i$ are again the NLS estimates of the errors defined in Equation (B.1).