Changing Business Dynamism and Productivity: Shocks vs. Responsiveness

Ryan A. Decker, John Haltiwanger, Ron S. Jarmin, and Javier Miranda*

September 2017

Abstract

The pace of business dynamism (as measured by indicators such as job reallocation) has declined in recent decades in the U.S., and theory suggests that this could have reduced productivity. At first glance the timing of changes in the pace of reallocation does not appear to match the changes in aggregate productivity: Productivity surged in the 1990s, led by the ICT sector, and has declined in the post-2000 period in ICT and more broadly, while overall reallocation and the entry rate of new firms declined throughout the 1980s, 1990s, and 2000s. However, pre-2000 trends in entry and reallocation largely reflected the productivity-enhancing consolidation of the retail trade sector, while the post-2000 period has seen declining entry and high-growth activity in the critical high tech sector. We study changing reallocation specifically, abstracting from changing startup rates by focusing within firm age groups where most of the variation in reallocation has occurred, and we give particular emphasis to the high tech sector given its outsized role in productivity fluctuations. We document changes in the way businesses respond to idiosyncratic productivity realizations, implying a quantifiable drag on productivity from a reduced pace of productivity-boosting reallocation. During the productivity slowdown of the post-2000 period, we find not only declining responsiveness but also rising within-industry productivity dispersion. Taken together, these findings are consistent with rising costs or declining incentives for factor adjustment in the U.S. and suggest that changing business dynamism is an important component of the U.S. productivity slowdown.

*Federal Reserve Board, University of Maryland and NBER, U.S. Census Bureau, and U.S. Census Bureau, respectively. We gratefully acknowledge financial support from the Kauffman Foundation. Cody Tuttle provided excellent research assistance. We thank John Abowd, Rudi Bachmann, Martin Baily, Jonathan Baker, Dave Byrne, Chris Foote, Lucia Foster, Clément Gorin, Bronwyn Hall, Matthias Kehrig, Pete Klenow, Kristin McCue, and conference or seminar participants at the 2015 Atlanta Fed Conference on Secular Changes in Labor Markets, the ASSA 2016 meetings, the 2016 Brookings "productivity puzzle" conference, the 3rd International ZEW conference, the 2016 ICC conference, Drexel University, the Federal Reserve Board, George Mason University, Georgetown University, the New York Fed, Princeton University, the University of Chicago, the spring 2017 Midwest Macro meetings, the 2017 UNC IDEA conference, and the 2017 NBER Summer Institute meetings for helpful comments. We are grateful for the use of the manufacturing productivity database developed in Foster, Grim and Haltiwanger (2016) as well as the revenue productivity database developed in Haltiwanger et al. (2016), and we thank Peter Schott for providing guidance and data for constructing import penetration data. Any opinions and conclusions expressed herein are those of the authors and do not necessarily represent the views of the U.S. Census Bureau or of the Board of Governors or its staff. All results have been reviewed to ensure that no confidential information is disclosed.

I. Introduction and motivation

Business dynamics—the process of business birth, growth, decline and exit—is a critical driver of the productivity-enhancing reallocative process that characterizes market economies. An optimal pace of business dynamics balances the benefits of productivity and economic growth against the costs associated with reallocation. While it is difficult to prescribe what the optimal pace should be, the pace of business dynamism in the U.S. has fallen over recent decades, and this downward trend accelerated after 2000.[1]

Hopenhayn and Rogerson (1993) show that dynamic frictions on factor (e.g., labor) adjustment reduce both job reallocation and productivity.[2] Thus, a *prima facie* concern arising from recent trends in business dynamism is that they may have reduced aggregate productivity growth. At first glance, recent fluctuations in U.S. productivity growth do not match up with patterns of business dynamism: productivity surged in the 1990s through the early 2000s before slowing after 2003 (Fernald (2014)), while aggregate startup activity and job reallocation fell throughout the 1980-2014 period. However, a more careful review of theory and evidence resolves the inconsistency: prior to 2000, the decline in entrepreneurship and reallocation was dominated by the productivity-boosting consolidation of the retail trade sector. [3] Business formation and job reallocation in the high tech sector actually rose during the 1980s and 1990s before declining after 2000 along with high-growth firm activity more generally (Haltiwanger, Hathaway, and Miranda (2014)), roughly coinciding with the ICT-driven surge in productivity from the late 1980s to early 2000s and subsequent decline after 2003 (Fernald (2014)).

We show that changes in how businesses respond to their idiosyncratic productivity conditions have been an important driver of both aggregate job reallocation and productivity in recent decades, especially in the high tech sector (we call this the "changing responsiveness" explanation). In high tech manufacturing, during the 1980s and 1990s plant-level survival and growth became more responsive to idiosyncratic TFP differences, while in the post-2000 period responsiveness declined substantially, particularly in terms of exit. Outside of high tech,

[1] See Davis et al. (2007), Haltiwanger, Jarmin and Miranda (2011), Reedy and Litan (2011), Hyatt and Spletzer (2013), Davis and Haltiwanger (2014) and Decker et al. (2014).
[2] The Hopenhayn and Rogerson (1993) finding is about the effect of an increase in adjustment frictions on the level of productivity, which implies reduced productivity growth if adjustment frictions increase repeatedly over time.
[3] Foster et al. (2006), Jarmin, Klimek and Miranda (2009), and others document the shift away from single unit establishment firms ("mom and pop shops") to national chains. Foster et al. (2006) and Foster et al. (2015) show that establishments of national chains are more productive and more stable. We discuss this evidence further below.

responsiveness declined throughout the 1980s-2000s. These findings are consistent with models of firm dynamics in which rising adjustment frictions reduce reallocation, increase the dispersion of marginal revenue products, and reduce productivity growth by making firms more cautious in responding to idiosyncratic productivity shocks.[4]

While we find support for the "changing responsiveness" explanation for changing reallocation rates, standard models also suggest another possible explanation that we call "changing shocks": since reallocation arises from the constantly shifting idiosyncratic productivity and profitability conditions faced by businesses, a decline in reallocation rates could arise from a decline in the dispersion or intensity of shocks. But we show that the dispersion of idiosyncratic productivity, both in and out of high tech, has risen steadily over time while the persistence of productivity exhibits little or no trend, ruling out the "changing shocks" explanation for aggregate reallocation patterns.

An alternative potential cause of changes in reallocation and the responsiveness of firms is the recent decline in startup rates. If young firms are more responsive to productivity shocks, changes in the average age of U.S. firms would mechanically reduce overall responsiveness, and the question of declining responsiveness and reallocation would boil down to the question of why startup activity has declined.[5] Consistent with earlier findings (Davis et al. (2006), Decker et al. (2014)), we show that the changing firm age structure induced by declining startup rates accounts for just one quarter of the overall decline in the job reallocation rate. We focus on changes in reallocation and responsiveness within firm age groups to mitigate the identification challenge posed by declining startup rates (while recognizing that the decline in startup rates is itself an important research area).[6]

We investigate these issues with a particular focus on the high tech component of the manufacturing sector where we can construct measures of establishment-level total factor productivity (TFP). These data allow us to evaluate both changes in business-level responses to their own productivity ("changing responsiveness") and changes in the distributional characteristics of productivity ("changing shocks"). We compare our results for high tech

[4] In addition to the canonical Hopenhayn and Rogerson (1993), see Cooper and Haltiwanger (2006), Cooper, Haltiwanger and Willis (2007, 2016), Elsby and Michaels (2013), and others.

[5] Young firms may be more volatile for a variety of reasons, such as the learning and uncertainty resolution dynamics hypothesized by Jovanovic (1982).

[6] For example, see Alon et al. (2017), who quantify the productivity implications of declining entry.

manufacturing to the remainder of the manufacturing sector, then we perform similar analyses using new data on firm-level labor productivity for the entire U.S. private sector. These exercises strongly suggest that the patterns we find in manufacturing generalize to other sectors.

Our results have implications for aggregate productivity: counterfactual exercises suggest that the increased responsiveness of the 1980s and 1990s yielded as much as half a log point annual boost in industry-level TFP in the high tech sector by the second half of the 1990s. The declining responsiveness of the 2000s yielded as much as a two-log-point drag on annual industry-level TFP by 2010. Moreover, evidence based on labor productivity suggests that the finding of declining responsiveness since 2000 generalizes beyond high tech manufacturing to other high tech businesses as well as other areas of the economy. The pre-2000 rise and post-2000 fall of productivity responsiveness in the high tech sector coincides with the ICT-driven rise and fall of aggregate productivity growth in the U.S.

The post-2000 decline in productivity responsiveness is robust and widespread across firm age groups and sectors, and the responsiveness decline in high tech manufacturing is also evident when we measure responsiveness in terms of equipment investment instead of employment growth. We find mixed evidence linking falling responsiveness with import competition. We also report the striking fact that within-industry labor productivity dispersion has risen since the late-1990s, consistent with rising adjustment frictions and inconsistent with slowing innovation in a Gort and Klepper (1982) framework.

Section II describes key facts about the declining pace of business dynamism. Section III describes the datasets we employ. In section IV, we use establishment-level data for manufacturing, with a particular focus on high tech, to study whether the evidence implies "changing shocks" or "changing responsiveness," and we analyze the implications of our findings for aggregate productivity growth. Section V looks beyond manufacturing and investigates the same questions using firm-level labor productivity and employment data for all U.S. sectors. Concluding remarks are in section VI.

II. Business Dynamics: Basic facts

A. Sectoral Patterns of Reallocation and Young Firm Activity

Starting with Davis et al. (2007), many studies have documented a decline in the pace of aggregate job reallocation and other indicators of business dynamism.[7] Decker et al. (2016) describe substantial cross-sector variation in patterns of reallocation: retail trade exhibits the strongest decline during the 1980s and 1990s, while information and high tech saw rising reallocation over that period before falling sharply after 2000. These patterns are depicted on Figure 1 (using HP trends) for selected NAICS sectors as well as high tech (as defined by Hecker (2005)).[8] As noted above, high tech is a particular focus of this paper due to its role in aggregate productivity dynamics (Fernald (2014)). Figure 2 illustrates similar patterns in the share of employment accounted for by young firms: retail trade saw declining startup activity throughout the 1980s-2010s, while information and high tech saw rising startup activity prior to 2000.[9]

Figures 1 and 2 also single out the high tech component of manufacturing. Reallocation and startup activity behave similarly in high tech manufacturing to the high tech sector more generally. Moreover, the information sector, which includes a heavy contingent of tech industries, behaves similarly to high tech broadly.

The changing patterns of young firm activity in Figure 2 account for some of the reallocation patterns in Figure 1. Figure 3 reports annualized changes in reallocation rates for select sectors (and economy wide) for three periods: the late 1980s to late 1990s (1987-1989 to 1997-1999), the late 1990s to mid-2000s (1997-1999 to 2004-2006), and the mid-2000s to early 2010s (2004-2006-2011-2013). We use three-year averages at business cycle peaks to abstract from cyclical concerns. For each sector, solid bars indicate the actual annualized change in reallocation rates over the period. The patterned (non-solid) bars indicate annualized changes resulting from a shift-share exercise that holds constant the age composition of businesses (at its initial state); that is, the non-solid bars describe within-age group changes in reallocation rates. We use seven age groups (firm age 0, 1, 2, 3, 4, 5, and 6+).

---

[7] Job reallocation is the sum of jobs created by expanding establishments and jobs destroyed by downsizing establishments, expressed as a rate by dividing by average two-year employment as in Davis, Haltiwanger, and Schuh (1996).

[8] Hecker (2005) defines industries as high tech based on the 14 four-digit NAICS industries with the largest share of STEM workers. The high tech sector thus defined includes industries in manufacturing (NAICS 3254, 3341, 3342, 3344, 3345, and 3364), information (5112, 5161, 5179, 5181, and 5182), and services (5413, 5415, and 5417). Notably, certain industries in the information sector are not high tech (e.g., book publishing).

[9] Guzman and Stern (2016) focus on extreme high-potential startups and also find 2000 to be an important turning point, with fewer high-growth outcomes for startups identified as having high potential.

During the 1990s (i.e., 1987-1989 to 1997-1999), the sharp decline in reallocation in retail trade and the increase in information are evident. The patterned bars show that falling young firm activity accounts for about a third of the reallocation decline in retail trade (i.e., two thirds of the decline occurred within age groups), and rising young firm activity accounts for about a tenth of the reallocation increase in information. The services sector saw a more modest reallocation decline during the 1990s which is entirely accounted for by falling young firm activity in that sector.

The pace of decline in several sectors accelerated after the late 1990s. This can be seen in services, which had a more modest decline during the early 1990s. More notably, though, reallocation rates in information fell markedly during the early 2000s after rising during the 1990s, with about a fifth of the early-2000s decline accounted for by falling young firm activity. Each sector continued declining during the late-2000s, and in each case the change in reallocation can be partially but not completely explained by falling young firm activity. This is the main inference we draw from Figure 3: while changing startup rates can account for a nontrivial portion of the overall change in job reallocation rates since the 1980s, most of the variation occurred within firm age groups. This finding encourages us to focus on changing patterns of responsiveness within firm age classes.

B. Possible Sources of Declining Startup Rates

A number of competing hypotheses may account for the variation in startup rates seen in recent decades. Changes in startup rates may endogenously reflect changes in the pace of innovation in an industry for reasons hypothesized by Gort and Klepper (1982): a period of rapid innovation leads to a surge in entry, reallocation and subsequent productivity growth.[10] Moreover, Gordon (2016) has argued that most of the 1980s-1990s high tech innovations had already been implemented by the early 2000s, and the productivity slowdown since that time is due to slowed innovation and implementation. Taken together, these hypotheses suggest that the changing pace of both startup activity and reallocation in the high tech sector in recent decades could have been caused by an exogenously changing pace of innovation.[11]

[10] Foster et al. (2017) provide supportive empirical evidence for these dynamics for the 1990s U.S. high tech sector.
[11] In the Gort and Klepper (1982) framework, declining innovation should be accompanied by declining dispersion of productivity within industries. While certainly not dispositive for the hypothesis of slowing innovation, we find the opposite below.

In retail trade, the share of sales and employment accounted for by single unit establishment firms fell from half to a third from 1977 to 2007 (see Foster et al. (2006), Jarmin et al. (2009), and Foster et al. (2016)). This dramatic transition is almost entirely accounted for by the rapid rise of large, national "big box" retailers, which are more productive (by about 30 log points) and have lower exit rates (by a factor of 15) than single-unit operations. Retail consolidations were likely facilitated by advances in information technology and globalization that permitted the development of large and efficient supply chains and distribution networks. Two key points are worth noting for our purposes: first, retail trade is an example of a sector in which declining reallocation and entrepreneurship has been productivity enhancing. Second, this change is reflected in the changing age structure of firms from which we abstract in our analyses.

Yet another factor contributing to the decline in startups is demographics-driven changes in labor force growth in the U.S. Karahan, Pugsley and Sahin (2015) show that variation in labor force growth driven by exogenous changes in population growth is positively associated with startup activity, an insight consistent with Hopenhayn (1992)-type models in which labor force growth is accommodated by adjustment in the number of firms.

Finally, the hypothesis that is the main focus of this paper—rising frictions inducing lower responsiveness of businesses—may be contributing to falling startup rates. An increase in frictions raises the cost of business activity and reduces the expected discounted value of profits for entrants, a key quantity governing entry in standard models. In this respect, our focus on within-age group variation may understate the contribution of declining responsiveness to the decline in productivity since the early 2000s.

III. Data and Measurement

The backbone dataset for our analysis is the Longitudinal Business Database (LBD), to which we attach other data as detailed below. The LBD includes annual location, employment and industry data for the universe of non-agriculture private sector establishments along with firm identifiers based on operational control (rather than an arbitrary taxpayer identifier).[12] We use the LBD for the period 1979-2013 (during which years consistent establishment-level NAICS codes are available from Fort and Klimek (2016)). Consistent with previous literature,

[12] See Jarmin and Miranda (2002) for a full description of the LBD.

we construct firm age based on the age of the firm's oldest establishment at the time the firm identifier first appears in the data, after which the firm ages naturally.

A. Manufacturing and TFP

We construct measures of TFP for over 2 million plant-year observations (1981-2010) using data from Foster, Grim, and Haltiwanger (2016) (hereafter FGH) combining the Annual Survey of Manufacturers (ASM) with the quinquennial Census of Manufacturers (CM). The ASM-CM is representative of the manufacturing sector in any given year, but it is a rotating sample so its longitudinal properties are inferior to those of the LBD. We therefore integrate the ASM/CM TFP data into the LBD to obtain establishment-level employment growth.[13]

We construct two alternative empirical measures of TFP for our analysis. The first, which has been commonly used in the literature (see, e.g., Baily, Hulten and Campbell (2001), Foster, Haltiwanger and Krizan (2001), Syverson (2011), Ilut, Kehrig and Schneider forthcoming), is a cost share-based index given by:

$$lnTFP_{et} = lnQ_{et}^{R} - \alpha_K lnK_{et} - \alpha_L lnL_{et} - \alpha_M lnM_{et} - \alpha_E lnE_{et} \qquad (1)$$

where Q^R is real output, K is real capital, L is labor input, M is materials, E is energy, α denotes factor elasticities, e denotes individual establishments and t denotes time. Output is total value of shipments plus the total change in the value of inventories, deflated using industry deflators from the NBER-CES Manufacturing Industry Database. Capital is measured separately for structures and equipment using a perpetual inventory method. Labor is total hours of production and non-production workers. Materials are measured separately for physical materials and energy where each is deflated by an industry-level deflator. Outputs and inputs are measured in constant 1997 dollars. Factor elasticities are estimated using industry-level cost shares (of total factor costs) with a Divisia index that allows cost shares to vary over time.[14] More details on measurement of output and inputs are in FGH.

This measure of TFP is a revenue-based measure and is increasingly referred to as TFPR. TFPR is defined by Foster, Haltiwanger and Syverson (2008) as P*TFPQ, where P is the plant-

[13] We use propensity score weights (based on a logit model on industry, firm size, and firm age) to adjust the ASM/CM/LBD sample to represent the LBD (in the cross section) in each year (see FGH for details). These weights are cross-sectionally representative in any given year but are not ideal for using samples of ASM/CM that are present in both t and t+1. We discuss this further below.

[14] Cost shares yield factor elasticities under the assumptions of cost minimization and full adjustment of factors. We are not assuming full adjustment for each plant at each unit of time but rather that this holds approximately when pooling across all plants in the same industry over time.

level price and TFPQ is the typical measure of plant-level technical efficiency in economic models such as the model we consider below. If plants are price takers, within-industry variation in TFPR only reflects TFPQ.[15] If plant-level prices are endogenous, TFPR still will be highly correlated with TFPQ in the adjustment cost framework we specify below. Moreover, as we show below, inferences using TFPR about changing responsiveness are still valid in such a framework. However, with endogenous prices, variation in dispersion in TFPR will not reflect just shocks to fundamentals such as TFPQ but also will endogenously reflect variation in adjustment costs.

Given possibly endogenous plant-level prices, we consider an alternative measure of productivity that has been increasingly used in the recent literature (see, e.g., Gopinath et. al. (2017) and Foster et. al. (2017)). To help illustrate this approach, consider a simple plant-level demand function $P_{et} = D_{et}Q_{et}^{\varphi-1}$ (where D_{et} is an idiosyncratic demand shock and $\varphi - 1$ is the inverse demand elasticity), a plant-level production function that is Cobb-Douglas with factor elasticity for factor *i* equal to α_i, and TFPQ equal to A_{et}. Then plant-level revenue is given by (lower case variables are in logs):

$$p_{et} + q_{et} = \beta_k k_{et} + \beta_l l_{et} + \beta_m m_{et} + \beta_e e_{et} + \varphi a_{et} + d_{et} \tag{2}$$

where $\beta_i = \varphi\alpha_i$ for factor *i*. That is, the β_i coefficients are the revenue elasticities that reflect both demand parameters and the production function factor elasticities. Given revenue elasticity estimates, the revenue productivity residual is given by:

$$RPR_{et} = \varphi a_{et} + d_{et}, \tag{3}$$

That is, RPR is solely a function of idiosyncratic TFPQ and demand. This implies that (as discussed in detail in Foster et al. (2017)) RPR can exhibit positive dispersion regardless of frictions and distortions. We estimate RPR by estimating the revenue function in (2) using the GMM approach in Wooldrige (2009) (see Appendix C for more details).

For each of these measures of productivity (which we denote as TFP for convenience), we take the log of TFP and deviate it from its detailed industry-by-year mean. These alternative measures are therefore within-industry measures that abstract from aggregate and industry-

[15] Assuming price taking behavior is not equivalent to assuming the homogenous goods and a single price in an industry. If plants within an industry have different product segments but are price takers within product segments then TFPR still only reflects fundamentals reflecting the potential quality differentials accounted for by price heterogeneity within an industry. TFPR is a referable measure to TFPQ in this case since it captures quality differentials.

specific shocks and are unaffected by mismeasurement of industry-level prices (Byrne and Corrado (2015, 2016)). We model TFP as an AR(1) process. The current-period realization of the idiosyncratic component of TFP is the shock, and we also consider innovations to these shocks by estimating the AR(1) process below.

In practice, we find that TFPR and RPR are highly correlated (about 0.8). This finding is consistent with the findings in Foster et al. (2017), and it is consistent with the findings in Foster, Haltiwanger and Syverson (2008, 2016) that TFPR and TFPQ are highly correlated (about 0.75) for the selected set of products where P and Q data are available to construct direct measures of TFPQ. Unsurprisingly, then, the main findings of our empirical analysis on changing responsiveness and changing shocks are robust to using TFPR or RPR. For the sake of brevity, we focus on the TFPR results in the main text, but we discuss the results for RPR throughout. In addition, the details of the results for RPR are provided in Appendix C.

B. Economywide Labor Productivity

In Section IV we extend our analysis to the entire economy by constructing measures of firm-level labor productivity. The RE-LBD combines LBD employment (collapsed from the establishment to the firm level) with revenue measures in the Census Bureau's Business Register (BR) (collapsed from the EIN to the firm level). Revenue data are available from 1996 to 2013; see Haltiwanger et al. (2016) for more details.[16] Consistent with previous literature, we construct annual firm employment growth rates on an "organic" basis to represent changes in establishment-level employment rather than artificial growth caused by mergers and acquisitions.

Similar to our TFP construction, we use (log) revenue per worker deviated from detailed (6-digit NAICS) industry-by-year means as a measure of firm labor productivity. We thereby control for price differences across industries such that our labor productivity measure is a within-industry relative gross output per worker measure; Foster, Haltiwanger and Krizan (2001, 2006) show that within-industry relative gross output per worker is highly correlated with within-industry relative value added per worker and strongly correlated with within-industry relative TFP (suggesting that materials and capital shares are similar across firms within

[16] About 20-percent of LBD firm-year observations cannot be matched to BR revenue data because firms can report income under EINs that may fall outside of the set of EINs that the Census considers part of that firm for employment purposes. We address potential match-driven selection bias by constructing inverse propensity score weights (separately for births, deaths, and continuers) such that the RE-LBD is representative of the LBD universe in terms of the size, age, employment growth rate, broad industry, and single/multi-unit structure of firms.

industries). We omit firms in the Finance, Insurance and Real Estate sectors (NAICS 52-53) from all analysis due to the difficulty of measuring output and productivity in those sectors.[17] As we show below, in our adjustment cost framework inferences regarding changing responsiveness can also be made using revenue per worker.

IV. Change in shocks vs. change in responsiveness

A. Theoretical motivation

Models of firm[18] dynamics suggest that a within-sector decline in the pace of reallocation is either due to a change in the volatility of shocks faced by firms or a change in firms' responses to those shocks. In appendix B we consider a kinked adjustment cost model consistent with canonical models of firm dynamics with adjustment frictions (a classic reference for our purposes is Hopenhayn and Rogerson (1993)). Firms face idiosyncratic productivity shocks, where the realization of productivity in the current period, A_{et}, is drawn from a persistent AR1 process. The decision rule for firms' net hiring rates is given by $g_{e,t} = f_t(A_{et}, E_{et-1})$, where the state variables are the productivity realization A_{et} and the initial employment E_{et-1}, both of which are observed prior to the growth decision.[19] We do not model entry or exit but discuss these margins below. For purposes of discussion in this section and appendix B, A_{et} is referred to as TFP or TFPQ. If there are demand shocks, this measure should be interpreted as a composite shock measure reflecting both TFPQ and demand shocks.

We calibrate the model and report numerical analysis to motivate the empirical specifications and moments we consider below (see appendix B for calibration details). Our model and calibration allows for endogenous plant level prices. Importantly for our empirical approach, revenue productivity measures with endogenous plant-level prices —TFPR or revenue labor productivity—are highly correlated with TFPQ (pairwise correlations of about 0.90) in a calibration with a plausible level of adjustment costs, though we do not target this moment in the calibration. Foster, Haltiwanger and Syverson (2008) find an empirical correlation of 0.75.

---

[17] Over 95 percent of firms, and over 99 percent of young (age less than five) firms, have only one establishment; for multi-unit firms we define each firm's industry based on its modal employment. In unreported exercises, we find our results are robust to a more sophisticated method of controlling for industry in multi-unit firms.

[18] We use the term "firms" loosely in this subsection for expositional ease.

[19] A similar rule would exist for investment in a model with capital. For net hiring rate dynamics, see, e.g., Cooper, Haltiwanger and Willis (2007, 2015) and Elsby and Michaels (2013). For investment dynamics, see, e.g. Cooper and Haltiwanger (2006).

We use the model to conduct two exercises. First, we investigate the "shocks" hypothesis, that is, the notion that falling reallocation reflects declining TFPQ dispersion. In our model, a decline in the dispersion of TFPQ, *ceteris paribus,* yields: (i) lower job reallocation; (ii) weaker responses of firm-level growth from t to $t+1$ to the realization of productivity in t (conditional on employment in t); and (iii) lower standard deviation of labor productivity.[20]

Second, we investigate the "responsiveness" hypothesis by varying the magnitude of adjustment frictions. An increase in adjustment frictions, *ceteris paribus,* yields: (i) lower job reallocation; (ii) weaker responses of firm-level employment growth from t to $t+1$ to the realization of productivity (using either TFPQ, TFPR or labor productivity) in t (conditional on employment in t); (iii) higher standard deviation of labor productivity; and (iv) lower Olley-Pakes (OP) covariance (i.e., the covariance between firm size and productivity) for both TFPQ and labor productivity, implying a decline in aggregate productivity consistent with a higher extent of misallocation. These model predictions are intuitive and are consistent with predictions in the literature.[21]

The Olley-Pakes (OP) covariance is based on a decomposition of the weighted mean of plant-level productivity. As we discuss in detail in Appendix B, a weighted mean only reflects industry-level productivity under the strong assumptions of CRTS and price taking behavior. With either decreasing returns or endogenous prices, industry-level productivity will reflect the variation in marginal revenue products with variation in the size across plants. We show that in spite of the difference between a weighted mean index of plant-level productivity and aggregate (industry-level) productivity, the OP covariance for either TFPQ or labor productivity will decrease with a rise in adjustment frictions. For labor productivity, this inference depends on starting from a benchmark with an empirically plausible level of adjustment frictions. As discussed in Appendix B, a frictionless benchmark is not empirically plausible since the pace of reallocation is implausibly large (over 100% of employment) in this case.[22]

---

[20] The second prediction stems in part from the inaction ranges that arise from our kinked adjustment cost design. As TFP dispersion falls there is a decrease in the fraction of firms that make zero adjustment (i.e., the "real options" effect). But declining TFP dispersion also implies smaller adjustments among those firms that do adjust (i.e., the "volatility" effect). Vavra (2014) argues that the volatility effect dominants the real options effect in the steady state, a general result extending back to Barro (1972).

[21] For example, Hopenhayn and Rogerson (1993) and Cooper, Haltiwanger and Willis (2007).

[22] This is based on a calibration using commonly used parameter values for technology, demand and shock processes. Bartelsman et. al. (2013) make a similar point about the use of OP covariances with labor productivity.

While the OP covariance is still an instructive moment for calibrating a structural model (as in Bartelsman et. al. (2013)), appropriate caution is needed in using weighted mean decompositions in quantifying the impact of changing frictions on aggregate productivity. In our empirical analysis below, we calculate a diff-in-diff counterfactual that we show (in Appendix B) tracks the impact of changes in adjustment frictions on aggregate (industry) productivity closely.

In considering the "responsiveness" hypothesis, we have in mind a broader interpretation than the simple adjustment frictions in our illustrative model. Given the recent literature on idiosyncratic distortions affecting the allocation of factors (e.g., Restuccia and Rogerson (2008), Hsieh and Klenow (2009) and Bartelsman et al. (2013)), an increase in the dispersion of distortions will have the same qualitative effects as an increase in frictions: lower productivity due to a weaker relationship of firm growth (and survival) with fundamentals. Moreover, while our simple model has only employment dynamics, we have in mind any type of increased friction that may impede adjusting the scale of operations at a firm.

Both declining TFPQ dispersion and rising adjustment frictions produce declining reallocation and declining Olley-Pakes covariance in our model calibration. To disentangle these forces we must capture the empirical evolution of both business-level responsiveness and the TFP distribution.[23] Moreover, the two hypotheses have opposite predictions for the dispersion of labor productivity: rising labor adjustment frictions would imply rising dispersion of labor productivity since increased frictions reduce the speed with which firms move their marginal revenue products toward equalization. In contrast, declining TFP dispersion would imply falling dispersion of labor productivity. A strength of our approach is that we seek identification by empirically studying multiple moments together.

Additional forces may be at work—beyond changes in shocks and frictions—that are not apparent from our illustrative model. In particular, the model neglects firm entry and exit, and as noted above there have been striking changes in entry dynamics in recent decades. Hopenhayn and Rogerson (1993) find that a rise in adjustment frictions will reduce entry and exit. In their model, the lower bound of productivity necessary for survival will decline with an increase in frictions. The empirical prediction, then, is that not only will firm growth for continuers become

[23] See Berger and Vavra (2017) for an application of the "shocks vs. responsiveness" approach in a different context; that paper and others cited therein likewise find an important role for the responsiveness factor in explaining aggregate outcomes.

less responsive to productivity, but so will exit. We explore this prediction in the empirical analysis below. Moreover, as discussed in the introduction, the firm dynamics of young firms differ from those of mature firms, so we control for potentially exogenous changes in entry rates by studying empirical moments within firm age groups.

We draw inference from theory by characterizing the evolution of key moments and reduced form relationships in the data. We do not identify a structural model of adjustment frictions, but we think this is a rich area for future research. For example, we do not take a stand on the exact form of adjustment costs (e.g., convex vs. non-convex), an area of interest in the literature. One potential use of our empirical findings, as suggested above, would be as moments to discipline such analysis.[24] A benefit of our reduced form approach is that it readily permits controlling for many different factors in a panel regression environment and allowing estimates to vary systematically by key firm characteristics such as detailed industry and firm age. In addition, we can use this reduced form approach to explore potential explanations for changes in the responsiveness to shocks that we detect.

B. Empirical Analysis of U.S. Manufacturing

In this section, we investigate these issues with establishment-level data for U.S. manufacturing with a particular focus on high tech manufacturing.[25] We first study the "shocks" hypothesis by exploring the evolution of TFP dispersion (i.e., the dispersion of establishment productivity draws), quantified as the standard deviation of (log) within-industry TFP (see Section III for TFP measurement details). Figure 4 reports TFP dispersion separately for plants of young and mature firms, in high tech and non-tech manufacturing.[26] We focus on low-frequency variation by reporting HP trends.

TFP dispersion has risen steadily in high tech manufacturing since the early 1980s and in non-tech manufacturing since the early 1990s. Within-industry TFP dispersion is large

[24] Cooper and Haltiwanger (2000) use regressions like ours in an indirect inference estimation of structural parameters of adjustment costs (capital adjustment, in their case). In their model, the marginal responsiveness of investment to profit shocks declines with increases in adjustment costs, either convex or non convex. Ilut, Kehrig and Schneider (forthcoming) estimate reduced-form policy functions with a focus on asymmetric responsiveness, finding that businesses respond more strongly to negative than to positive shocks. These authors do not study changes in this asymmetry over time, a potentially interesting question given our findings.

[25] These include NAICS codes 3341 (computer and peripheral equipment), 3342 (communications equipment), 3344 (semiconductor and other electronic components), 3345 (navigational, measuring, electromedical, and control instruments), 3254 (pharmaceutical and medicine), and 3364 (aerospace product and parts).

[26] Our unit of analysis in this section is the establishment (plant), but the LBD permits us to classify plants based on the age of the firm to which they belong.

(consistent with, *e.g.*, Syverson (2004, 2011)); for example, a level of 0.4 (40 log points) on Figure 4 implies that a plant one standard deviation above the mean for its industry is about $e^{0.4} \approx 1.5$ times as productive as the mean. Notably, within-industry TFP dispersion is about the same for plants of young and mature firms. Figure C1 in appendix C shows very similar results for the alternative RPR productivity measure based on Wooldridge (2009).

Plant dynamics depend not only on dispersion but also on persistence of idiosyncratic TFP: plants facing adjustment costs are more likely to respond to TFP shocks if TFP is more persistent (Cooper and Haltiwanger (2006); Cooper, Haltiwanger and Willis (2007)). Our data are not ideally suited for estimating TFP persistence and innovations, but Figure A4 in appendix A suggests that persistence is reasonably stable with an estimated AR(1) coefficient of about 0.6 to 0.7, and trends of TFP innovation dispersion (Figure A5) match trends of TFP dispersion.

Figures 4 and A4 suggest that changing reallocation is not driven by changing TFP dispersion or persistence. Consider high tech: Figure 1 shows reallocation rising during the 1990s then falling after 2000. For dispersion and persistence of TFP to account for the reallocation trend we would expect dispersion and/or persistence to mimic these patterns; or, conversely, given the patterns of TFP dispersion and persistence, we should see rising reallocation in the manufacturing sector in the post-2000 period. That we see the opposite is strong evidence against the "shocks" hypothesis for declining reallocation during that period.

We now turn to the "responsiveness" hypothesis by directly estimating the relationship between productivity and growth (and survival) at the establishment level. Our main outcome variable of interest is establishment employment growth from year t to $t+1$ using the Davis, Haltiwanger and Schuh (1996) (hereafter DHS) concept that accommodates entry and exit (by using the two-year average of employment as the denominator). We estimate the following:

$$\begin{aligned} g_{e,t+1} = \sum_{age=y,m} & [\beta_{age} TFP_{et} + \delta_{1age} TFP_{et} * Trend_t \\ & + \delta_{2age} TFP_{et} * Trend_t^2\,] * I_{age,et} \quad + X'_{et}\Theta + \varepsilon_{e,t+1} \end{aligned} \tag{4}$$

where $g_{e,t+1}$ is the DHS employment growth rate for establishment e from time t to $t+1$, TFP_{et} is (log) industry-deviated TFP for establishment e at time t, and $Trend_t$ is a simple linear time trend. The responsiveness to TFP in terms of the main and trend effects are permitting to vary by firm age with $I_{age,et}$ a dummy variable for young (age<5, subscript y) and mature (subscript m) plants (these dummy variables are denoted Young and Mature in the discussion below). X_{et} is a set of controls discussed below including year effects to capture national trends or cyclical shocks. We estimate equation (3) using our propensity score weights relating the ASM/CM to the LBD.

While this is a reduced form specification, it is broadly consistent with the specifications of selection and growth dynamics from the literature discussed above. First, it is consistent with the model calibration exercises in Appendix B that show, estimating the equivalent of equation (4) on simulated data, that an increase in adjustment frictions reduces the responsiveness of plant-level growth to lagged realizations of TFP. Second, by using DHS growth rates, we can incorporate both the extensive margin (exit) and the intensive margin of plant-level growth. Standard empirical specifications of exit in the literature (see, e.g., Syverson (2011)) relate the decision to exit between t and $t+1$ to the realization of TFP in period t along with other controls (*e.g.*, endogenous state variables such as size, which is part of our X_{et} as described below). As we have already noted, adjustment cost models of employment growth yield predictions that relate the growth in employment from time t to $t+1$ to the realization of TFP in t along with period-t size. In this sense, equation (4) produces a reduced-form yet direct estimate of policy functions generated by canonical models.[27]

Our question is whether the response to idiosyncratic productivity shocks has changed over time. The inclusion of the $Trend_t$ variable allows us to estimate a time-varying relationship between productivity and growth. In unreported results we have considered alternative ways to capture a changing trend (*e.g.*, interacting a linear trend with decade dummies), and results are robust to considering such alternatives. In exercises described below, we exploit this time-varying productivity responsiveness estimation to study the changing contribution of job reallocation to aggregate productivity growth.

We estimate specification (4) for 1981-2010 with the following controls in X_{et} (in addition to year effects): establishment size, firm size, state effects and a state-level business cycle indicator (the change in state-level unemployment rate).[28] We also interact the state-level

[27] TFP in period t is measured for calendar year t while establishment growth is measured from March of t to March of t+1. Thus, the empirical timing of the data is closer to the timing in the theoretical specifications in Appendix B than might first appear. In Appendix B, we show declining responsiveness of firm-growth to current or lagged realizations in productivity from an increase in adjustment frictions.

[28] For firm size effects, we use firm size classes in period *t*. For establishment size effects, we have considered both establishment size classes and log employment at the establishment level in period *t*. We obtain very similar results for both cases, and in the paper we use log employment at the establishment level.

cyclical indicator with plant-level TFP following FGH. The cyclical variables are all interacted with the young and mature dummies; our liberal inclusion of cyclical indicators is intended in part to avoid contamination of results from the Great Recession and other business cycle factors.

On Table 1, we report the main effects for TFP by firm age group and the interactions with the trend terms. Columns 1 and 2 show the growth regressions of equation (4); columns 3 and 4 show a linear probability model with exit as the dependent variable (but otherwise identical to equation (4)). The estimates for β_y and β_m are given by the "TFP*Young" and "TFP*Mature" rows. These positive (negative) coefficients show that, consistent with previous literature, productivity and growth (exit) are positively (negatively) related.[29] The growth coefficients are stronger for establishments of young firms, consistent with intense selection working on recently started businesses; the exit coefficients follow the same pattern in non-tech manufacturing, though interestingly this is not the case in high tech. The positive (negative) relationship between productivity and growth (exit) at the establishment level is consistent with a positive contribution of reallocation to aggregate productivity growth.

The estimates of δ_{1y} and δ_{1m} are given by the "TFP*Young*Trend" and "TFP*Mature*Trend" rows of Table 1, respectively. These coefficients show how the marginal responsiveness of establishments to their idiosyncratic productivity has changed with time. Notably, in high tech manufacturing δ_{1y} and δ_{1m} are positive (negative) and significant for the growth (exit) of plants of both young and mature firms, with the exception of the exit coefficient for mature firms, suggesting that productivity responsiveness generally strengthened in the early years of the sample (which begins in 1980), while the coefficients are close to zero among non-tech establishments. Both inside and outside of high tech, however, the growth (exit) coefficients on the quadratic term (δ_{2y} and δ_{2m}) are negative (positive).

We next graphically illustrate the implications of the combined linear and quadratic trend terms. Since TFP is measured relative to industry-year means, we can calculate the growth differential between a "productive" plant—the plant with TFP one standard deviation above its industry mean—and the average plant in an industry by multiplying the total regression coefficient (including trend effects) by the within-industry TFP standard deviation.[30] To abstract

[29] The coefficients relating productivity with growth or exit are statistically significant at the 1 percent level in all but one case: the coefficient for exit among young high tech establishments is significant at the 10 percent level.
[30] We set the cyclical indicator (state change in unemployment) to zero to evaluate effects at a neutral cyclical state.

from changing TFP dispersion, we fix the standard deviation at 0.41 for high tech and 0.37 for non tech (roughly the respective averages across time). Figure 5 shows the resulting growth differentials averaged by decade.

First note that young firm plants are more responsive to productivity than are mature firm plants, especially in high tech: in the 1980s (black bars), the growth differential among young high tech plants was 14 percentage points (meaning the plant with productivity one standard deviation above its industry mean grew 14 percentage points faster, over a one year period, than the plant with industry mean productivity), compared with 6 percentage points among mature high tech plants. This and Figure 4 imply that the high pace of reallocation of young-firm plants is not driven by a high variance of TFP but rather by a high responsiveness to TFP differences consistent with, for example, a learning model. The difference in responsiveness between plants in young and mature firms implies that overall responsiveness depends in part on the age distribution—hence our within-age group approach.

Our main focus is the variation in responsiveness over time. First, consider high tech manufacturing. For plants in young firms, the growth differential rises from 14 to 16 percentage points from the 1980s to the 1990s then declines to 9 percentage points in the 2000s. For plants in mature firms, responsiveness initially declines modestly from the 1980s to the 1990s then accelerates into the 2000s, with the growth differential stepping down from 6 to 5 percentage points then dropping to 3 percentage points.

Next, consider the non-tech results. Again, plants in younger firms are more responsive to TFP. Among young firms, the growth rate differential was about 10 percentage points in the 1980s, 9 percentage points in the 1990s, and 6 percentage points in the 2000s. Among mature firms, the growth differential was about 5 percentage points throughout the time period.

On Figure C3 of appendix C, we report exercises using the Wooldridge (2009) RPR as our TFP measure.[31] In high tech manufacturing, young-firm RPR responsiveness rises then falls as with TFPR and mature-firm RPR responsiveness falls sharply in the post-2000 period in a manner similar to the TFPR results. For non-tech plants, the overall drop in responsiveness from the 1980s to the 2000s is similar with RPR compared to TFPR although more of the drop in

[31] The decline from the 1990s to the 2000s among young high tech businesses is not as notable in the RPR-based regressions as it is in the TFPR-based regressions, but it is still significant as we show below in aggregate productivity counterfactuals.

responsiveness happens during the early part of the period (versus the later part for TFPR). The RPR results tell roughly the same story as the TFPR results, with reasonable similarity both quantitatively and qualitatively.

Figure A1 in appendix A shows charts for the exit coefficients reported on columns 3 and 4 of Table 1 and discussed above. Part of the growth responsiveness pattern is driven by selection dynamics associated with changing exit responsiveness. Among young-firm high tech establishments, exit selection intensified from the 1980s to the 1990s then weakened in the 2000s; young non-tech establishments saw steadily weakening selection throughout the period. Among mature-firm plants, selection intensity did not vary notably until it weakened somewhat in the 2000s. The findings on exit are interesting in their own right as they imply that in the post-2000 period low-productivity plants are more likely to survive, constraining aggregate productivity (and potentially raising TFP dispersion).

An alternative story is that rising dispersion of TFP (and its innovations) in the post-2000 period not only could be partially endogenous to changing selection but also could independently contribute to weakening growth responsiveness: in the presence of non-convex adjustment costs, higher TFP dispersion widens inaction bands and reduces the frequency of adjustment. But these concerns are not likely to be playing a dominant role: during the 1990s, we find increased responsiveness of exit in high tech despite mild increases in TFP dispersion, a finding that also holds for RPR (Figure C1 in appendix C). More broadly, the combined 30-year patterns of dispersion and responsiveness, across age and industry groups, cannot tell the alternative story coherently. As noted above, our model and a broader literature theorize that the "frequency" effect of widening inaction bands is not likely to dominate the "volatility" effect of larger adjustment-conditional changes in employment due to higher TFP dispersion.

Taken together, these results have important implications for the evolution of firm dynamics in recent decades. The way in which individual businesses respond to their idiosyncratic realizations of productivity has changed over time. The positive relationship between realized productivity and subsequent employment growth remains robust, but it has weakened (particularly since 2000). Through the lens of firm dynamics models, our results can be viewed as evidence that establishment-level policy functions have changed over time, particularly for young businesses but also for older ones.[32] In the post-2000 period, these changes are consistent with an increase in adjustment costs or other frictions that reduce marginal responsiveness to productivity in these models. The changes are most striking among high tech businesses, where we observe a pattern of rising and falling productivity responsiveness that coincides with the ICT-driven acceleration and deceleration of aggregate productivity growth documented by Fernald (2014) and others.

C. Implications for aggregate (industry-level) productivity

How important are the changes in responsiveness for aggregate fluctuations in productivity? For this purpose, we compute the following diff-in-diff counterfactual:

$$\Delta_t^{t+1} = \sum_e (\theta_{e,t+1}^T - \theta_{e,t+1}^{NT}) a_{et} \tag{5}$$

where θ_{et+1}^T is the predicted employment share for establishment e in period $t+1$ based upon the full model that includes trend patterns in responsiveness (the T superscript refers to "trend"), and θ_{et+1}^{NT} is the predicted employment share for establishment e in period $t+1$ predicted by the estimated model with parameters reflecting responsiveness at the beginning of the sample period (that is, we set the trend terms δ_{ij} equal to zero, so NT means "no trend").[33] The employment share prediction for an establishment in a given period ($\theta_{e,t+1}^T$) is based on the actual realizations of productivity and initial employment for that establishment in the previous period, fed through the estimated growth rate model; $\sum_e \theta_{e,t+1}^T a_{et}$ is the predicted aggregate productivity in period $t+1$ holding all establishments' productivity constant at period-t levels and only changing establishments' employment shares.[34]

[32] These results contrast with Karahan, Pugsley and Sahin (2016), who argue that the dynamics of incumbent firms (in a Hopenhayn framework) have not changed over this same time period. These authors point to aggregated average growth rates for various incumbent age classes as evidence for stable incumbent firm dynamics. We differ from their approach by directly estimating incumbent firm policy functions at the establishment level. Viewed through their framework, our results suggest that factors in addition to changes in the growth of the labor force are likely relevant for understanding the decline in startup rates, though the labor force demographic evidence is an important component of the final explanation.

[33] We set the cyclical effects to zero by setting the state-level change in unemployment to zero.

[34] This approach is related to, but distinct from, the accounting productivity decompositions in the literature (see, e.g., Foster, Haltiwanger and Krizan (2001) for a review). Our present approach differs since it focuses only on model-driven reallocation that is identified to be the reallocation arising specifically from variation in productivity across businesses. In addition, this approach focuses only on the reallocation components since the exercise holds constant the productivity distribution at the micro level between t and t+1. Decker et al. (2017) use the Dynamic Olley-Pakes (DOP) decomposition developed by Melitz and Polanec (2015) to show that these accounting decompositions also imply a decline in the contribution of the change in the covariance terms (often interpreted as indicators of allocative efficiency) in the post-2000 period. Alon et al. (2017) likewise use the DOP decomposition

In Appendix B we show that this diff-in-diff counterfactual closely tracks aggregate productivity effects of changing responsiveness in our benchmark adjustment cost model. Figure B6 shows that the diff-in-diff counterfactuals for both TFP and labor productivity track the decline in aggregate productivity from an increase in adjustment costs. Figure B6 also shows that the OP covariance declines with an increase in adjustment costs but, given the assumed endogenous prices in the benchmark calibrated model, the OP covariance declines more rapidly with adjustment costs than does aggregate productivity or the diff-in-diff counterfactuals. This is because the OP decomposition is based on a decomposition of the weighted mean of micro productivity. The latter matches aggregate (industry-level) productivity with CRTS and perfect competition. With decreasing returns or product differentiation, the weighted mean and aggregate productivity exhibit highly correlated changes with increases in adjustment costs but the weighted mean and the OP covariance decline more rapidly.

Another attractive feature of this diff-in-diff counterfactual is that it only captures the effect of time-varying responsiveness within firm age groups. Differences in responsiveness between young and mature firms will be present in both the counterfactual with and the counterfactual without the trend, as will the changing age structure of firms overall. Moreover, this diff-in-diff design mechanically abstracts from potential effects of changing TFP dispersion.

We report Δ_t^{t+1} for each year on Figure 6. For example, the observation for $t+1=1981$ has $\Delta_{1980}^{1981}=0$ because the trend variable begins then, and for high tech the year 2001 again gives $\Delta_{2000}^{2001}=0$. But the 2004 observation for high tech shows that if responsiveness from 2003 to 2004 had been at the 1981 pace instead of the actual pace (as estimated by our model) then the productivity index in 2004 would have been about half a log point higher ($\Delta_{2003}^{2004}=-0.005$). For high tech manufacturing plants, the increasing responsiveness over the 1980s and 1990s yields an implied counterfactual increase in the index that peaks at about half a log point per year in the 1990s. The sharp decline in responsiveness during the post-2000 period implies a decline in the productivity index of as much as 2 log points per year by 2010. Some caution needs to be used in interpreting the magnitude at the end points—and certainly extrapolating out of sample—since the pattern in Figure 6 is driven by fitting a quadratic trend. But we regard our findings as

but focus on the cumulative contribution of changes in entry rates. As discussed in the text and appendix B, the accounting decompositions have shortcomings for quantifying productivity effects in the presence of decreasing returns or product differentiation.

implying that the drag on this index of industry level productivity due to the decline in responsiveness may be quite substantial.

Figure C4 in appendix C reports the same exercise but using the RPR productivity concept; in high tech, the RPR results are quite similar—qualitatively and quantitatively—to the TFPR results from Figure 6, while outside of high tech the productivity drag implied by the RPR regressions starts somewhat sooner. The basic message of the TFPR and RPR results is the same, however, particularly in high tech: changing responsiveness has quantitatively large implications for aggregate productivity. In high tech, changing responsiveness starts to be a drag on productivity around 2003, about the time that Fernald (2014) finds a trend break in productivity growth in the IT sector. Outside of high tech, both the TFPR and RPR results show a decline in aggregate productivity from the 1980s to the 2000s from declining responsiveness.

Some caution should be used in interpreting our counterfactual results as yielding patterns that mimic actual aggregate (industry-level) productivity growth since there may be changes in the within-plant productivity components of aggregate (industry-level) growth that we have not estimated in this context. Fernald (2014), Byrne et al. (2016) and Gordon (2016) highlight many factors that are likely contributing to within-plant (and within-firm) declines in productivity growth in the post-2000 period. In addition to the factors they emphasize, there may be a role for declining entrepreneurship in declining within-firm productivity growth given the contribution of young firms to innovative activity (Acemoglu et al. (2013) and Alon et al. (2017)). We examine the within-firm productivity growth patterns below.

D. Changing Business Models

A possible explanation for changing responsiveness is that it reflects changes in business models that are benign for productivity. For example, perhaps businesses increasingly respond to shocks by adjusting their capital stock instead of labor (a sort of capital/labor substitution). We repeat the regressions from equation (4), replacing the employment growth rate with the investment rate (investment divided by initial capital) and adding initial capital as an additional control.[35] This regression therefore includes the key state variables: productivity, initial employment, and initial capital stock.

Table 2 reports the regression results for high tech manufacturing, and Figure 7 shows

[35] See appendix D for more detail. Note that the intuition linking adjustment costs to employment growth applies equally to investment in models of firm dynamics (Cooper and Haltiwanger (2006)).

results analogously to Figure 5. As with employment, young firms' investment is more responsive than mature firms. Investment responsiveness in high tech manufacturing displays a qualitatively similar pattern to employment responsiveness, with a significant decline in the 2000s among young firms: in the 1990s, a young-firm plant with TFP one standard deviation above its industry-year mean had an equipment investment rate 8 percentage points higher than the plant at the mean; this differential is about 3 percentage points in the post-2000 period. The decline in employment responsiveness was not accompanied by stronger investment responsiveness in high tech manufacturing. Among non-tech manufacturing businesses, however, there is rising investment responsiveness from the 1980s to the 1990s, with responsiveness remaining elevated in the 2000s, suggesting that capital-labor substitution may play a modest role outside of the high tech sector. More broadly, we cannot rule out other forms of capital investment—like intangibles—as substitute adjustment mechanisms.

In appendix D, we describe two other exercises exploring changes in business model. First, we find mixed evidence that industries facing increased import competition saw bigger declines in responsiveness, suggesting that globalization may be an interesting avenue for future work. Second, we find no evidence that industry composition shifts within high tech manufacturing explain declining responsiveness.

V. Beyond Manufacturing

Thus far we have focused on the manufacturing sector in which we have high-quality TFP data. An important question is whether the patterns of productivity dispersion and responsiveness we have described are present outside manufacturing. For example, the information sector has been a key contributor to U.S. innovation in recent years. Moreover, changes in startup rates and in the dispersion and skewness of firm growth rates are even more dramatic in non-manufacturing components of the high tech sector (Decker et al. (2016)).

We next conduct the same exercises as in section IV, but with firm-level gross output per worker as our productivity concept and with the full private sector as our sample.[36] Output per worker cannot be used to directly track the pattern of shocks, but our adjustment cost framework

[36] This exercise requires that we assign each firm an industry code; we do this by choosing the industry that accounts for the largest share of the firm's employment. Our results are robust to an alternative approach to controlling for firm industry activity, which is not surprising given the prevalence of single-unit firms.

in appendix B shows that moments based on output per worker move systematically with changes in adjustment frictions. We employ RE-LBD data (described in Section III), which permit the measurement of revenue per worker at the firm level for the entire U.S. private, non-farm sector from the mid-1990s to 2013.[37] An independent contribution of this section is new evidence on the relationship between productivity, firm-level growth, and reallocation dynamics outside manufacturing.[38]

The inferences we draw in this section recognize that output per worker endogenously reflects not only TFP but also changes in adjustment frictions. However, as our benchmark adjustment cost model illustrates, several empirical moments based on output per worker are informative for changing adjustment frictions. First, recall from section IV that an increase in adjustment frictions implies an increase in the within-industry dispersion of labor productivity: adjustment frictions dampen the tendency for marginal revenue products to be equalized, implying higher labor productivity dispersion. Second, an increase in adjustment frictions also reduces the responsiveness of firm-level employment growth from t to $t+1$ to the realization of revenue labor productivity in t (controlling for employment in t). Finally, an increase in adjustment frictions reduces the diff-in-diff counterfactual using labor productivity in a manner that tracks the implications for aggregate productivity.

A. Productivity and growth at the firm level

Figure 8 reports the standard deviation of within-industry labor productivity for young and mature firms, in and out of high tech; labor productivity dispersion has risen for each of these groups (note that our definition of high tech now includes certain industry groups in services and information as well as manufacturing).[39] Notably, unlike TFP, labor productivity is more dispersed among young than mature firms; younger firms likely face greater learning or other frictions and may also be more heterogeneous in capital intensity.

Our finding of rising within-industry productivity dispersion is broadly consistent with other work documenting increased differences between firms. For example, Andrews, Criscuolo and Gal (2015) find a widening productivity gap between "frontier firms" and others, arguing

[37] We omit finance, insurance, and real estate (NAICS 52-53) from our sample.
[38] Relatively little is known about these issues outside manufacturing. Exceptions include several retail trade studies (Foster, Haltiwanger and Krizan (2006), Jarmin, Klimek and Miranda (2009) and Foster et al. (2015)).
[39] Bils, Klenow and Ruane (2017) argue that rising revenue productivity dispersion in U.S. manufacturing datasets reflects sampling-based mismeasurement; our results here, however, show rising revenue productivity dispersion even in administrative data.

that the pace of technological diffusion has slowed. While the diffusion hypothesis could play a role, our estimates of TFP persistence (appendix A Figure A4) suggest that the group of "frontier firms" is sufficiently fluid to somewhat limit the diffusion story's explanatory power. Weakening responsiveness of growth and survival to productivity is an alternative, but not mutually exclusive, explanation. Both explanations allow for a decoupling of technological progress and aggregate productivity growth.[40] Declining responsiveness is consistent with increased adjustment frictions impeding marginal revenue product equalization.

Rising labor productivity dispersion is also evidence against the "shocks" hypothesis for falling reallocation in various U.S. sectors. If only shocks were increasing we should expect to see rising labor productivity dispersion and rising reallocation. Instead, the latter is declining. We next turn to the "responsiveness" hypothesis. We estimate equation (4)—the regression we used to measure changing TFP responsiveness in manufacturing—except that we now use firm-level data (vs. establishment), labor productivity in place of TFP, all U.S. sectors (except finance, insurance and real estate), and only the years 1997-2013.[41]

Table 3 reports results of these regressions. The first two columns report regressions using the DHS growth rate denominator inclusive of exit; the last two columns report results using only a binary exit outcome as the dependent variable. Figure 9 graphically shows the time series pattern of the growth coefficients; as with TFP results, we report the growth rate differential between the firm one standard deviation above its industry mean and the mean. Growth is indeed related to revenue labor productivity, as theorized; that is, firms with higher output per worker are more likely to grow.[42] Figure A2 in appendix A shows a strong relationship between labor productivity and exit as well. Young firms are particularly sensitive to labor productivity, including on the exit margin, indicating that labor productivity is correlated with selection mechanisms. Moreover, the relationship of labor productivity with growth and survival has weakened over time, particularly among young high tech firms (where the growth

[40] Andrews, Criscuolo and Gal (2015) (ACG) provide evidence of rising productivity dispersion within broad sectors using ORBIS data on both labor productivity (similar to our approach here) and multifactor productivity (similar to our analysis in Section IV). ACG measure the difference between "frontier firms" and average firms, where the frontier firms are usually defined as the top 50 or 100 firms within a broad (2-digit) sector, and in the case of the U.S. their unit of analysis is actually the establishment (Pinto Ribeiro, Menghinello and De Backer (2010)).

[41] We also apply propensity score weights to address the fact that the RE-LBD covers 80 percent of LBD businesses; see Section III for RE-LBD details.

[42] Growth differentials for labor productivity may seem large compared with TFP-based differentials from the previous section; this is partly because labor productivity dispersion is higher than TFP dispersion.

differential has fallen by 10 percentage points), consistent with the TFP-based evidence from Section IV. Broadly speaking, the evidence suggests that the survival and growth differential between high- and low-productivity firms is declining over time, particularly in high tech.

The data on both labor productivity dispersion and the relationship linking labor productivity with growth and survival indicate that the TFP-based patterns we found in manufacturing are likely to hold in other sectors. Again the framework of firm dynamics models, applied to our evidence, suggests that slowing reallocation is a symptom of increased frictions rather than changes in the distribution of idiosyncratic productivity shocks.

A. Reallocation and aggregate labor productivity

Following the approach from Section IV, we quantify the labor productivity regression results by relating them to aggregate productivity growth by using the diff-in-diff counterfactual approach from equation (5). As shown in Figure B.6, in the calibrated model the diff-in-diff counterfactual using labor productivity tracks the impact of increased adjustment frictions on aggregate (industry-level) productivity quite well.[43]

The results for implementing the diff-in-diff counterfactual for the high tech (not just manufacturing) and non-tech sectors is presented in Figure 10. By 2013, the weakening responsiveness of growth and survival to productivity accounts for more than 5 log points in the diff-in-diff counterfactual. This implies that if responsiveness had stayed at the 1996 levels that aggregate productivity in 2013 would be 5 log points higher. In contrast to the TFP-based results from manufacturing, our labor productivity-based calculations for the entire economy show a similar pattern for firms inside and outside high tech. In unreported results we find that this is driven by particularly strong declines in the sensitivity of exit to productivity among firms outside high tech; moreover, within manufacturing specifically, we do find stronger results in high tech than outside of it in the labor productivity counterfactuals, as in the TFP counterfactuals.

B. Changing Patterns of Within-Firm Productivity Growth

Has a falling productivity contribution from reallocation been offset by stronger within-firm productivity growth? We use the firm labor productivity database to construct two related

[43] Even though we are using the diff-in-diff counterfactual, it is useful to note that for gross output per worker that the weighted mean of micro productivity tracks gross output per worker at the industry level quite well. For example, Figure A3 shows that the industry-level weighted mean indices yield patterns very similar to the traditional index of gross output per worker using BLS industry-level indices.

but distinct measures of within-firm productivity growth. The first measure is the simple unweighted mean of annual within-firm productivity growth. The second is the employment-weighted mean of annual within-firm productivity growth using time-t employment weights for productivity growth from t to $t + 1$. We compute these measures at the 6-digit industry level then aggregate using time-invariant employment weights for each industry.[44]

Figure 11 shows trends in within-firm productivity growth, both weighted and unweighted, for the average industry, separately for high tech and non tech. Note the following: First, for high tech, within-firm productivity growth declines using both measures. Second, for non tech, weighted within-firm productivity growth declines but the unweighted measure exhibits less systematic variation. Third, the weighted measure is much larger than the unweighted measure for both tech and non tech, and the unweighted measure is always negative for non tech and turns negative for high tech early in the sample. This third finding might be surprising since it implies negative productivity growth for the average firm. However, as discussed by Decker et al. (2017), the unweighted measure overwhelmingly describes very small firms (more than 90 percent of firms have fewer than 20 employees). Decker et al. (2017) further show that the positive difference between the weighted and the unweighted mean reflects a positive relationship between within-firm productivity growth and initial shares (i.e., larger firms have higher within-firm productivity growth).

In sum, within-firm improvements (e.g., innovation by incumbents) have not quickened to compensate for weaker reallocation.[45]

VI. Conclusion

Reallocation has declined in all sectors—particularly the high tech sector—since the early 2000s, following more mixed patterns in the 1980s and 1990s. Within-industry TFP dispersion has risen in recent decades, both in and out of high tech, which cannot explain the patterns of reallocation over the same time. Rather, changing reallocation reflects changes in the marginal response of businesses to idiosyncratic productivity conditions, consistent with

[44] Importantly, these measures of within-firm productivity growth exploit the RE-LBD's longitudinal links and are therefore distinct from exercises that measure average productivity growth among specific groups of firms.

[45] Alon et al. (2017) also use the Dynamic Olley Pakes decomposition method described by Melitz and Polanec (2015) to study the productivity slowdown. The authors show that declining entry has had a significant cumulative negative effect on aggregate productivity growth, consistent with our emphasis on the contribution of changing firm dynamics.

increased frictions on labor adjustment. Our counterfactual exercises suggest that changing business dynamism can account for a significant drag on aggregate productivity, as much as 2 log points in high tech manufacturing and more than 5 log points economy-wide in recent years. The timing of reallocation and responsiveness patterns in high tech manufacturing is consistent with the timing of the productivity slowdown, which evidence indicates was driven by ICT-producing and using industries. Importantly, our evidence abstracts from the confounding effect of declining startup rates since we study responsiveness within firm age groups.

Our main results are based on plant-level TFP in high tech manufacturing, but the results extend to manufacturing industries more broadly and are robust to TFP measurement methodology. Further, even outside manufacturing we find rising labor productivity dispersion and a weakening of the relationship between firm-level growth and labor productivity.

In addition to shedding light on the drivers of declining business dynamism, these findings comprise a novel contribution to the literature on the U.S. productivity slowdown in the post-2000 period. Our results are not simply driven by declining startup rates; moreover, productivity dispersion within industries has risen in the post-2000 period, while a slowing pace of innovation would produce falling dispersion in a Gort and Klepper (1982) framework. Slowing reallocation and business-level responsiveness is therefore an important component of the productivity slowdown and, at least, is complementary to innovation-based explanations such as Gordon (2016) (and is consistent with Byrne, Fernald and Reinsdorf (2016) and Syverson (2016), who find that the productivity slowdown is not an artifact of mismeasurement).

We document several other interesting patterns. The responsiveness of equipment investment to TFP in high tech manufacturing follows a similar pattern to employment responsiveness, rising during the 1980s and 1990s then falling sharply after 2000, while investment responsiveness in non-tech manufacturing was flat throughout the 1990s and 2000s after rising in the 1980s. The strong relationship between growth and productivity that has previously been documented for TFP in manufacturing also holds for labor productivity in other sectors. The decline in productivity-enhancing reallocation has not been offset by stronger within-firm labor productivity growth.

We do not here study specific policy or other factors that may be contributing to declining responsiveness, a large task that we leave for future work. Our theoretical framework focuses on increased adjustment costs, but broader interpretations may be plausible.

Globalization may have played a role in subdued business-level growth responsiveness by facilitating cross-border factor adjustment (see appendix D). Other explanations could include any forces that raise the cost of, or incentive for, factor adjustment; possibilities include unlawful discharge regulations, occupational licensing rules, scope of intellectual property protection, land use regulations, rules or norms that increase job match specificity, or various other state or federal regulations.[46] Declining intensity of competition or increased prevalence of winner-take-all economics could also produce some of the empirical patterns we document (as argued by De Loecker and Eeckhout (2017)). Given the implications of declining responsiveness for productivity growth, this is an important area for future research.

[46] Using industry variation, Goldschlag and Tabarrok (2014) find no evidence that federal regulation counts relate with changes in the pace of gross flows, but both state and federal regulation may present further scope for research. Davis and Haltiwanger (2014) find evidence relating employment protection policies to lower rates of reallocation, consistent with earlier work by Martin and Scarpetta (2012) and Autor, Kerr and Kugler (2007). Kleiner and Krueger (2013) review occupational licensing data and research. Molloy et al. (2016) find that states with tighter land use restrictions as of the early 2000s did not see larger declines in labor flows in recent decades, but the effect of changes in land use regulations is unknown; Hsieh and Moretti (2017) estimate significant static misallocation from land use regulations.

References

Alon, Titan M., David Berger, Rob Dent, and Benjamin Pugsley. 2017. "Older and Slower: The Startup Deficit's Lasting Effects on Aggregate Productivity Growth." *Mimeo.*

Acemoglu, Daron, Ufuk Akcigit, Nicholas Bloom, and William R. Kerr. 2013. "Innovation, Reallocation, and Growth." NBER Working Paper No. 18993.

Andrews, Dan, Chiara Criscuolo, and Peter N. Gal. 2015. "Frontier Firms, Technology Diffusion and Public Policy: Micro Evidence from OECD Countries." OECD: *The Future of Productivity: Main Background Papers.*

Autor, David H., William R. Kerr, and Adriana D. Kugler. 2007. "Does Employment Protection Reduce Productivity? Evidence from US States." *The Economic Journal* 117 no. 521:F189-F217.

Bachmann, Rüdiger and Christian Bayer. 2013. "Wait-and-See Business Cycles." *Journal of Monetary Economics* 60(6): 704-19.

Baily, Martin Neil, Charles Hulten, and David Campbell. 1992. "Productivity Dynamics in Manufacturing Plants." In *Brookings Papers on Economic Activity: Microeconomics*, ed. Clifford Winston and Martin Neil Baily. Washington, DC: Brookings Institution Press.

Baily, Martin Neil, Eric Bartelsman, and John Haltiwanger. 2001. "Labor Productivity: Structural Change and Cyclical Dynamics." *Review of Economics and Statistics*, 83(3), 420-433.

Barro, Robert J. 1972. "A Theory of Monopolistic Price Adjustment." *The Review of Economic Studies* 39(1): 17-26.

Becker, Randy A., John Haltiwanger, Ron S. Jarmin, Shawn D. Klimek, and Daniel J. Wilson. 2006. "Micro and Macro Data Integration: The Case of Capital." Chap. 12 in *Studies in Income and Wealth* edited by Dale W. Jorgenson, J. Steven Landefeld, and William D. Nordhaus, vol. 66, *A New Architecture for the U.S. National Accounts*, pp. 541-609. Chicago, IL: National Bureau of Economic Research, University of Chicago Press.

Bernard, Andrew B., J. Bradford Jensen, and Peter K. Schott. 2006. Survival of the Best Fit: Exposure to Low-Wage Countries and the (Uneven) Growth of US Manufacturing Plants. J*ournal of International Economics* 68(1): 219-237.

Bils, Mark, Peter J. Klenow, and Cian Ruane. 2017. "Misallocation or Mismeasurement?" Working paper.

Bloom, Nicholas. 2009. "The Impact of Uncertainty Shocks." *Econometrica*, 623-85.

Byrne, David M. 2015. "Domestic Electronics Manufacturing: Medical, Military, and Aerospace Equipment and What We Don't Know about High-Tech Productivity," FEDS Notes 2015-06-02. Board of Governors of the Federal Reserve System.

Byrne, David M. and Carol A. Corrado. 2015. "Recent Trends in Communications Equipment Prices." *FEDS Notes*, no. 2015-06-02.

Byrne, David M. and Carol A. Corrado. 2016. "ICT Prices and ICT Services: What Do They Tell Us about Productivity and Technology?" Working paper.

Byrne, David M., John G. Fernald and Marshall B. Reinsdorf. 2016. "Does the United Have a Productivity Slowdown or a Measurement Problem?" *Brookings Papers on Economic Activity*, Spring 2016.

Cairó, Isabel. 2013. "The Slowdown in Business Employment Dynamics: The Role of Changing Skill Demands." *Mimeo.* At http://www.econ.upf.edu/eng/graduates /gpem/jm/cairo.html.

Cooper, Russell and John Haltiwanger. 2000. "On the Nature of Capital Adjustment Costs." NBER Working Paper No. 7925.

Cooper, Russell and John Haltiwanger. 2006. "On the Nature of Capital Adjustment Costs." *Review of Economic Studies* 73(3); 611-33.

Cooper, Russell, John Haltiwanger, and Jonathan Willis. 2007. "Search Frictions: Matching Aggregate and Establishment Observations." *Journal of Monetary Economics* 54 Supplement 1: 56-78.

Cooper, Russell, John Haltiwanger, and Jonathan Willis. 2015. "Dynamics of labor demand: Evidence from plant-level observations and aggregate implications," *Research in Economics,* 69: 39-50.

Davis, Steven J. and John Haltiwanger. 1991. “Wage Dispersion Between and Within U.S. Manufacturing Plants, 1963-1986.” *Brookings Papers on Economic Activity: Microeconomics*, 1991, pp. 115-200.

Davis, Steven J. and John Haltiwanger. 2014. “Labor Market Fluidity and Economic Performance.” NBER Working Paper 20479 (published in the 2014 Federal Reserve Bank of Kansas City Jackson Hole Conference Symposium).

Davis, Steven J., John Haltiwanger, Ron Jarmin, and Javier Miranda. 2007. "Volatility and Dispersion in Business Growth Rates: Publicly Traded versus Privately Held Firms." Chap. 2 in *NBER Macroeconomics Annual 2006* edited by Daron Acemoglu, Kenneth Rogoff, and Michael Woodford, vol. 21, pp. 107-80. Cambridge, MA: MIT Press.

Davis, Steven J., John C. Haltiwanger, and Scott Schuh. 1996. *Job Creation and Destruction.* Cambridge, MA: MIT Press.

De Loecker, Jan and Jan Eeckhout. 2017. “The Rise of Market Power and the Macroeconomic Implications.” NBER Working Paper no. 23687.

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

Decker, Ryan A., John Haltiwanger, Ron S. Jarmin, and Javier Miranda. 2016. “Where Has All the Skewness Gone? The Decline in High-Growth (Young) Firms in the U.S.” *European Economic Review* 86 (July) 4-23.

Decker, Ryan A., John Haltiwanger, Ron S. Jarmin, and Javier Miranda. 2017. “Declining Dynamism, Allocative Efficiency, and the Productivity Slowdown” *American Economic Review Papers and Proceedings* (May).

Dunne, Timothy, Lucia Foster, John Haltiwanger, and Kenneth R. Troske. 2004. “Wage and Productivity Dispersion in United States Manufacturing: The Role of Computer Investment.” *Journal of Labor Economics* 22 no. 2:397-430.

Elsby, Michael W. L. and Ryan Michaels. 2013. “Marginal Jobs, Heterogeneous Firms, and Unemployment Flows.” *American Economic Journal: Macroeconomics* 5 no. 1:1-48.

Ericson, Richard, and Ariel Pakes. 1995. “Markov-Perfect Industry Dynamics: A Framework for Empirical Work.” *Review of Economic Studies* 62(1): 53–82.

Fernald, John. 2014. "Productivity and Potential Output Before, During, and After the Great Recession." Chap. 1 in *NBER Macroeconomics Annual 2014* edited by Jonathan A. Parker and Michael Woodford, vol. 29, pp. 1-51. Cambridge, MA: MIT Press.

Fort, Teresa and Shawn Klimek. 2016. "The Effects of Industry Classification Changes on US employment composition." Mimeo.

Foster, Lucia, Cheryl Grim, and John Haltiwanger. 2016. "Reallocation in the Great Recession: Cleansing or Not?" *Journal of Labor Economics.*

Foster, Lucia, Cheryl Grim, John Haltiwanger, and Zoltan Wolf. 2017. "Innovation, Productivity Growth and Productivity Dispersion," mimeo.

Foster, Lucia, John Haltiwanger, Shawn Klimek, C.J. Krizan and Scott Ohlmacher. 2015. "The Evolution of National Chains: How we got here", Center for Economic Studies Working Paper No. CES-WP-15-10.

Foster, Lucia, John Haltiwanger, and C.J. Krizan. 2001. "Aggregate Productivity Growth: Lessons from Microeconomic Evidence." In *New Developments in Productivity Analysis*, eds. Edward Dean, Michael Harper and Charles Hulten, 303-363. Chicago: University of Chicago Press.

Foster, Lucia, John Haltiwanger, and C.J. Krizan. 2006. "Market Selection, Reallocation and Restructuring in the U.S. Retail Trade Sector in the 1990s." *The Review of Economics and Statistics* 88(4): 748-58.

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

Foster, Lucia, John Haltiwanger, and Chad Syverson, (2016), "The Slow Growth of New Plants: Learning about Demand?" Economica, 83: 91–129.

Goldschlag, Nathan and Alex Tabarrok. 2014. "Is Regulation to Blame for the Decline in American Entrepreneurship?" GMU Working Paper in Economics No. 15-11.

Gordon, Robert J. 2016. *The Rise and Fall of American Growth: The U.S. Standard of Living since the Civil War*. Princeton: Princeton University Press.

Guzman, Jorge and Scott Stern. 2016. "The State of American Entrepreneurship: New Estimates of the Quantity and Quality of Entrepreneurship for 15 US States, 1988-2014." *Mimeo*.

Haltiwanger, John. 2016. “Firm Dynamics and Productivity: TFPQ, TFPR and Demand Side Factors." *Economia*, Forthcoming.

Haltiwanger, John, Ian Hathaway, and Javier Miranda. 2014. “Declining Business Dynamism in the U.S. High-Technology Sector.” The Kauffman Foundation.

Haltiwanger, John, Ron S. Jarmin, Robert Kulick and Javier Miranda. 2016. In *Measuring Entrepreneurial Businesses: Current Knowledge and Challenges* edited by John Haltiwanger, Erik Hurst, Javier Miranda, and Antoinette Schoar. Chicago: University of Chicago Press.

Haltiwanger, John, Jarmin, Ron, and Javier Miranda. 2009. “Business Dynamics Statistics: An Overview.” Kauffman Foundation: Other Research. At http://www.kauffman.org/uploadedFiles/BDS_handout_011209.pdf.

Haltiwanger, John, Ron Jarmin, and Javier Miranda. 2011. “Historically Large Decline in Job Creation from Startups and Existing Firms in the 2008-09 Recession.” Business Dynamics Statistics Briefing No. 5, The Kauffman Foundation.

Haltiwanger, John, Ron S. Jarmin, and Javier Miranda. 2013. “Who Creates Jobs? Small vs. Large vs. Young.” *Review of Economics and Statistics* 95(2): 347-61.

Haltiwanger, John, Robert Kulick, and Chad Syverson, 2016. “Misallocation Measures: Glowing Like the Metal on the Edge of a Knife.” Mimeo.

Hecker, Daniel. 2005. “High-technology employment: a NAICS-based update,” *Monthly Labor Review* (U.S. Dept. of Labor and U.S. Bureau of Labor Statistics), Volume 128, Number 7, July 2005: 58.

Hugo A. Hopenhayn. 1992. “Entry, Exit, and firm Dynamics in Long Run Equilibrium.” *Econometrica*, Vol. 60(5): 1127-50.

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

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

Hsieh, Chang-Tai and Peter J. Klenow. 2017. “The Reallocation Myth.” 2017 Jackson Hole paper.

Hsieh, Chang-Tai and Enrico Moretti. 2017. “Housing Constraints and Spatial Misallocation.” Mimeo.

Hyatt, Henry R. and James Spletzer. 2013. “The Recent Decline in Employment Dynamics.” *IZA Journal of Labor Economics* 2(3): 1-21.

Ilut, Cosmin L., Matthias Kehrig, and Martin Schneider. Forthcoming. “Slow to Hire, Quick to Fire: Employment Dynamics with Asymmetric Responses to News.” *Journal of Political Economy*.

Jarmin, Ron S., Shawn Klimek, and Javier Miranda. 2009. "The Role of Retail Chains: National, Regional and Industry Results." Chap. 6 in *Studies in Income and Wealth* edited by Timothy J. Dunne, J. Bradford Jensen, and Mark J. Roberts, vol. 68, *Producer Dynamics: New Evidence from Micro Data*, pp. 237-62. Chicago, IL: National Bureau of Economic Research, University of Chicago Press.

Jarmin, Ron S., and Javier Miranda. 2002. "The Longitudinal Business Database." CES Working Paper no. 02-17. Washington, D.C.: Center for Economic Studies, U.S. Census Bureau.

Jovanovic, Boyan. 1982. “Selection and the evolution of industry.” *Econometrica*, 50(3): 649–70.

Karahan, Fatih, Benjamin Pugsley, and Aysegul Sahin. 2016. “Understanding the 30-Year Decline in the Startup Rate: A General Equilibrium Approach.” Mimeo.

Kleiner, Morris and Alan B. Krueger. 2013. “Analyzing the Extent and Influence of Occupational Licensing on the Labor Market.” *Journal of Labor Economics* 31 no. 2:S173-S202.

Levinsohn, James and Amil Petrin (2003). “Estimating Production Functions Using Inputs to Control for Unobservables.” *Review of Economic Studies* 70 no. 2:317-41.

Lucas, Robert E. 1978. “On the Size Distribution of Business Firms.” *The Bell Journal of Economics* 9 no. 2:508-23.

Martin, John P. and Stefano Scarpetta. 2012. “Setting It Right: Employment Protection, Labour Reallocation and Productivity.” *De Economist* 160 no. 2:89-116.

Melitz, M. J., & Polanec, S. (2015). Dynamic Olley-Pakes productivity decomposition with entry and exit. *The RAND Journal of Economics*, *46*(2), 362–375.

Molloy, Raven, Christopher L. Smith, Riccardo Trezzi, and Abigail Wozniak. 2016. “Understanding Declining Fluidity in the U.S. Labor Market.” *Brookings Papers on Economic Activity*.

Olley, G. Steven and Ariel Pakes. 1996. "The Dynamics of Productivity in the Telecommunications Industry." *Econometrica*, 64 no. 6:1263-98.

Pierce, Justin and Peter Schott. 2016. "The Surprisingly Swift Decline of U.S. Manufacturing Employment." *American Economic Review* 107 no. 7:1632-62.

Pinto Ribeiro, Samuel, Stefano Menghinello and Koen De Backer. 2010. "The OECD ORBIS Database: Responding to the Need for Firm-Level Micro-Data in the OECD." *OECD Statistics Working Papers* No. 2010/01, OECD Publishing, Paris.

Reedy, E.J. and Robert E. Litan. 2011. "Starting Smaller; Staying Smaller: America's Slow Leak in Job Creation." Kauffman Foundation Research Series: Firm Foundation and Economic Growth. At http://www.kauffman.org/research-and-policy/starting-smaller-staying-smaller-americas-slow-leak-in-job-creation.aspx.

Schott, Peter. 2008. "The Relative Sophistication of Chinese Imports." *Economic Policy*, 5-49.

Syverson, Chad. 2004. "Product Substitutability and Productivity Dispersion." *Review of Economics and Statistics* 86(2): 534-50.

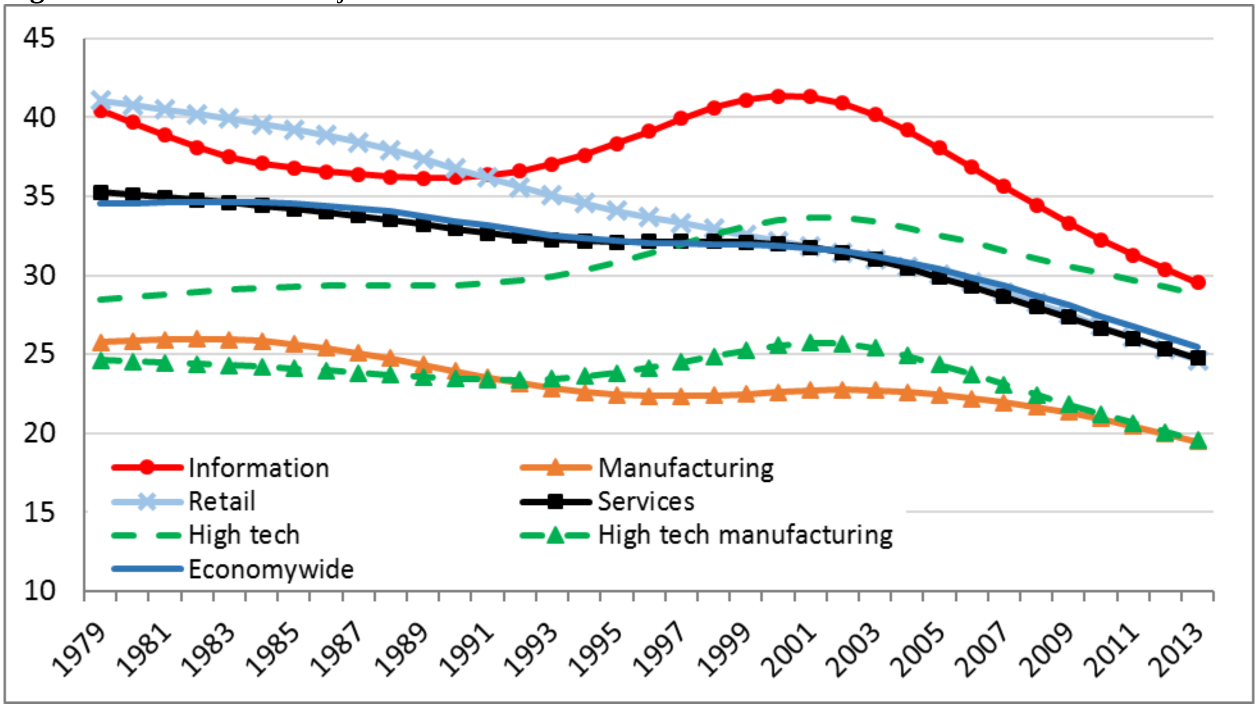
Syverson, Chad. 2011. "What Determines Productivity?" *Journal of Economic Literature* 49(2): 326-65.

Syverson, Chad. 2016. "Challenges to Mismeasurement Explanations for the U.S. Productivity Slowdown." NBER Working Paper no. 21974.

Vavra, Joseph. 2014. "Inflation Dynamics and Time-Varying Volatility: New Evidence and SS Interpretation." *Quarterly Journal of Economics* 129(1): 215-58.

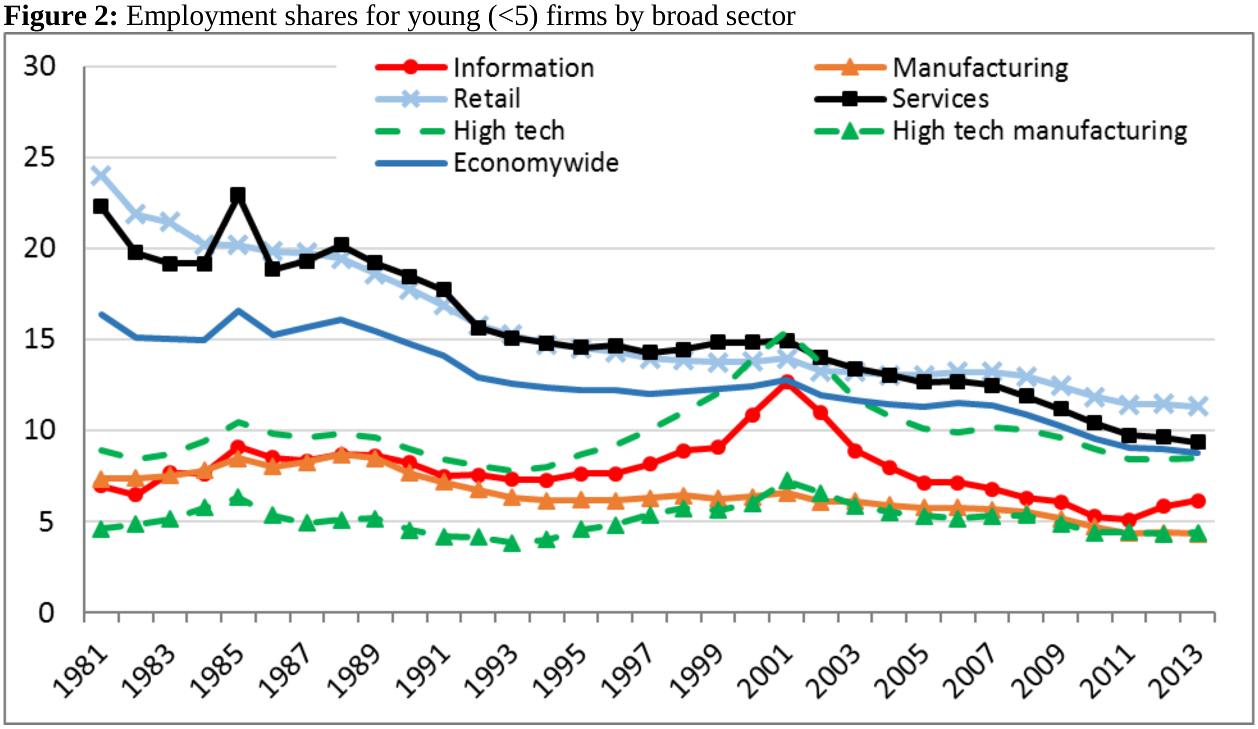
Wooldridge, Jeffrey M. 2009. "On Estimating Firm-Level Productivity Functions Using Proxy Variables to Control for Unobservables." *Economics Letters* 104 no. 3:112-4.

Figure 1: Sectoral trends in job reallocation



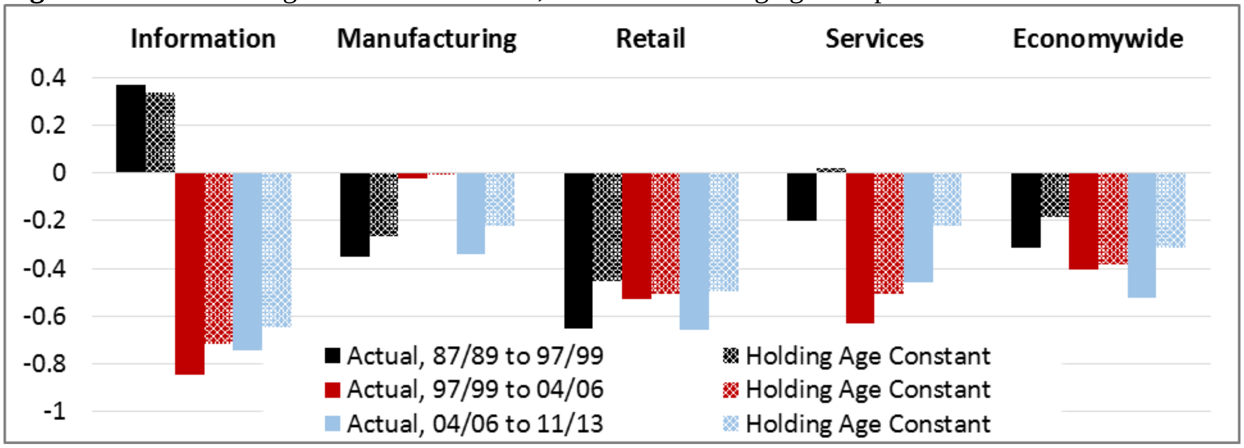
Note: Y axis does not start at zero. HP trends using parameter set to 100. Industries defined on a consistent NAICS basis; high tech is defined as in Hecker (2005). Data include all firms (new entrants, continuers, and exiters). Author calculations from the Longitudinal Business Database (LBD).

Figure 2: Employment shares for young (<5) firms by broad sector



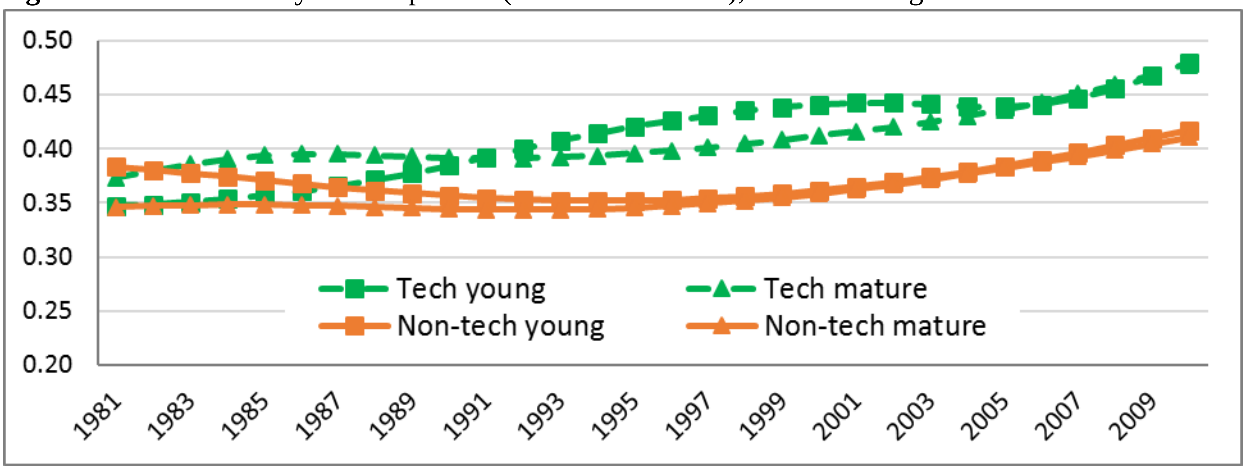
Note: Young firms have age less than 5. Industries are defined on a consistent NAICS basis; high tech is defined as in Hecker (2005). Data include all firms (new entrants, exiters, and continuers). Author calculations from the LBD.

Figure 3: Annual change in reallocation rate, actual and holding age composition constant



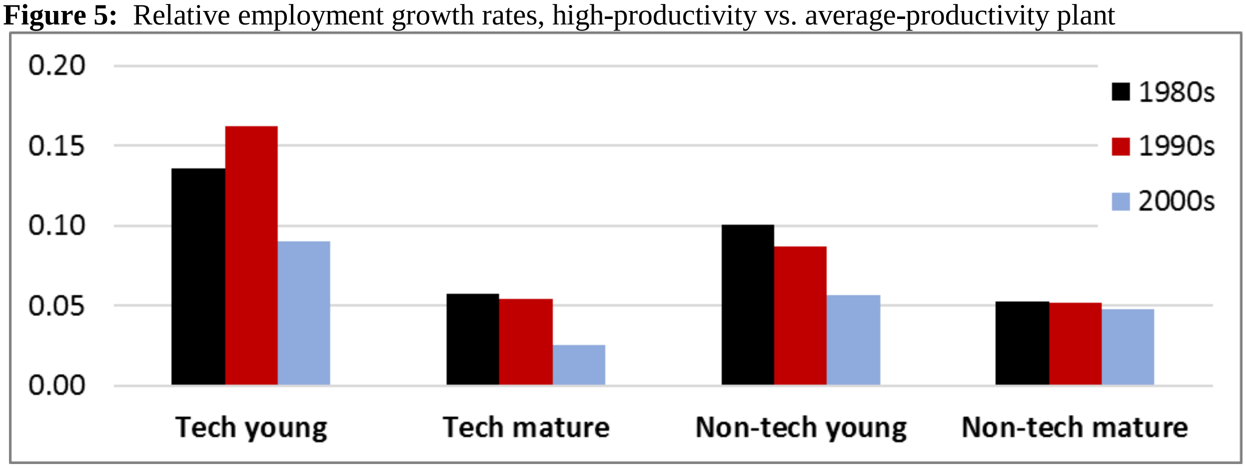
Note: Sectors are defined on a consistent NAICS basis. Author calculations from the LBD.

Figure 4: Within-industry TFP dispersion (standard deviation), manufacturing



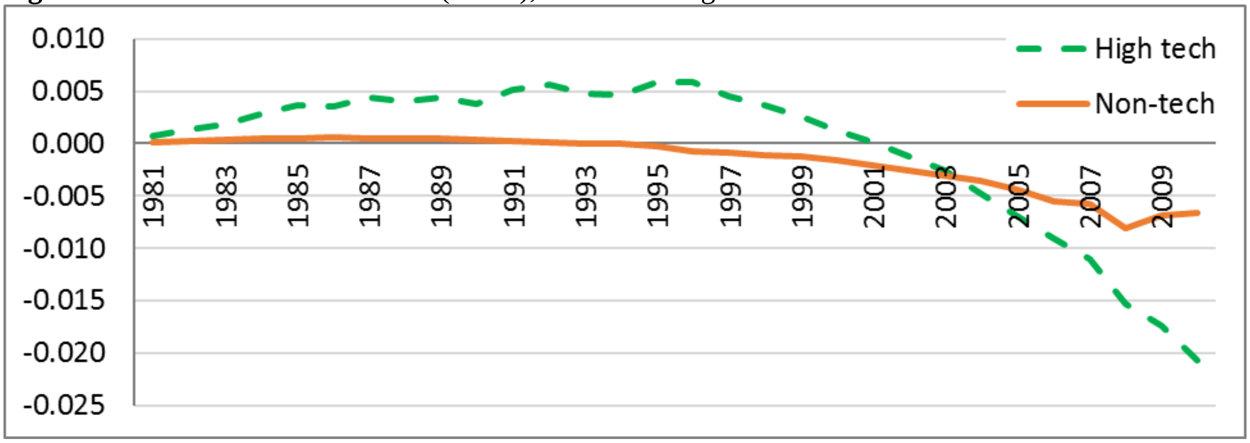
Note: Y axis does not start at zero. Young firms have age less than 5. Standard deviation of within-detailed industry log TFPR. High tech defined as in Hecker (2005). Author calculations from the LBD, the Annual Survey of Manufacturers (ASM), and the Census of Manufacturers (CM). HP Trends.

Figure 5: Relative employment growth rates, high-productivity vs. average-productivity plant



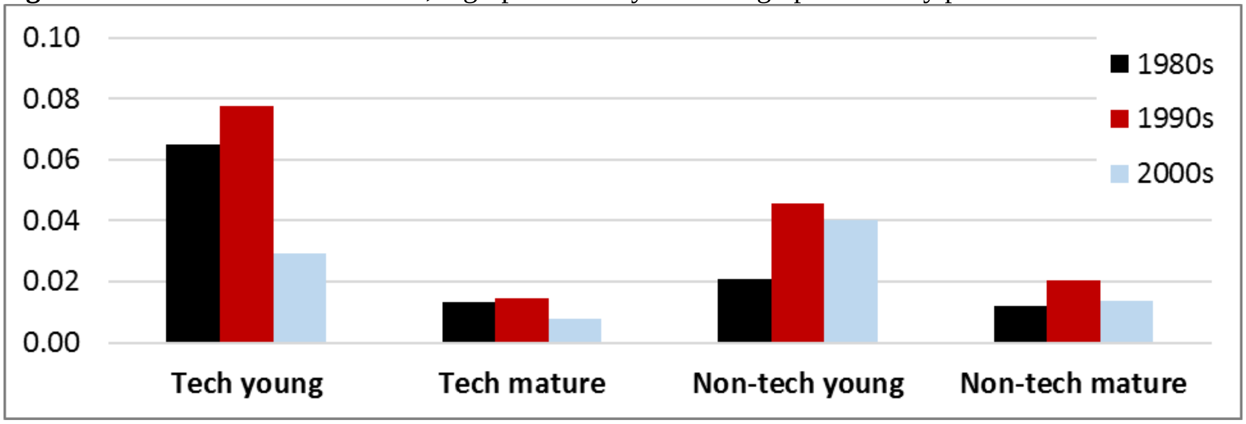
Note: Young firms have age less than 5. High tech is defined as in Hecker (2005). Growth rate of plant with TFP one std. dev. above industry mean vs. industry mean. Author calculations from the LBD, the ASM, and the CM.

Figure 6: Diff-in-diff counterfactual (TFPR), manufacturing



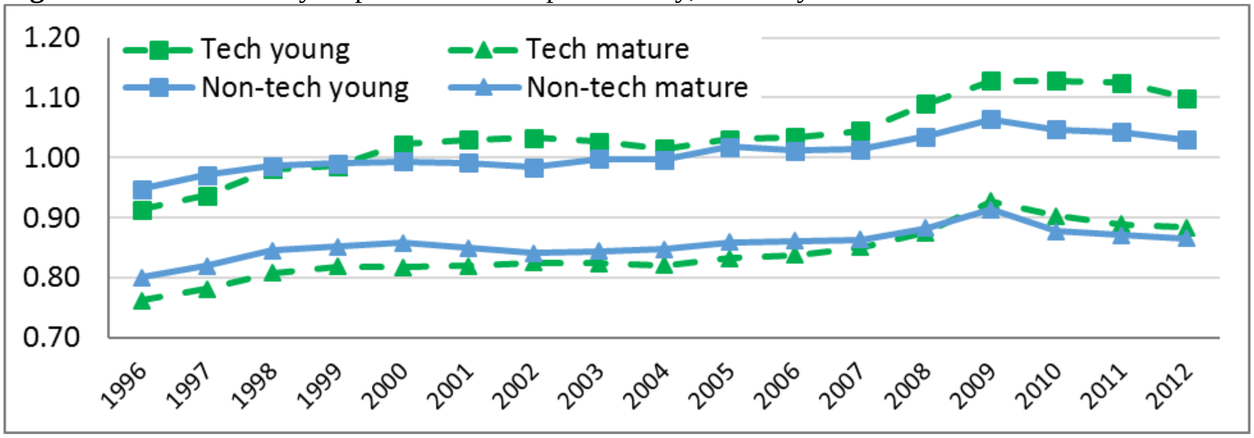
Note: Figure depicts diff-in-diff counterfactual as described in the text. High tech is defined as in Hecker (2005). Author calculations from the LBD, the ASM, and the CM.

Figure 7: Relative investment rates, high-productivity vs. average-productivity plant



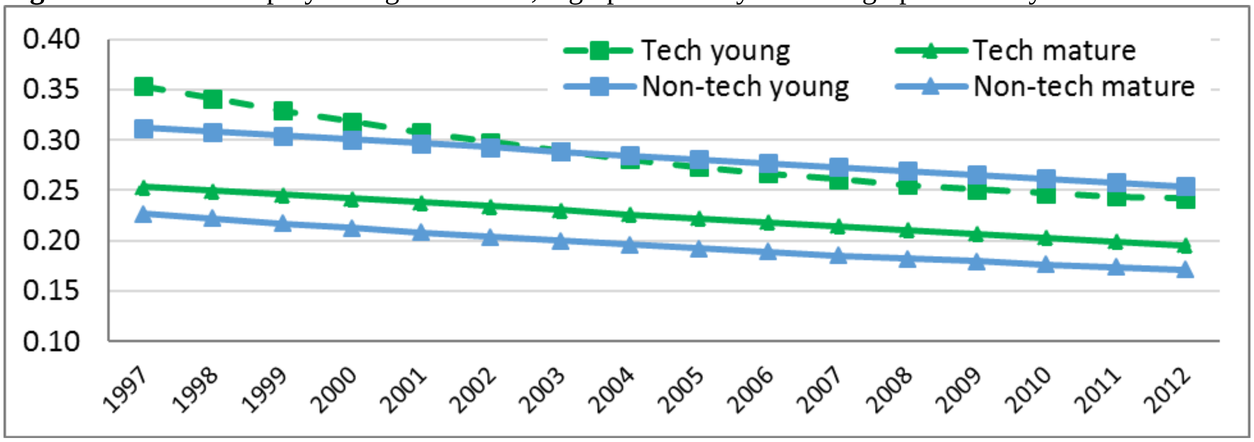
Note: Young firms have age less than 5. High tech is defined as in Hecker (2005). Investment rate of plant with TFP one std. dev. above industry mean vs. industry mean. Author calculations from the LBD, the ASM, and the CM.

Figure 8: Within-industry dispersion in labor productivity, economywide



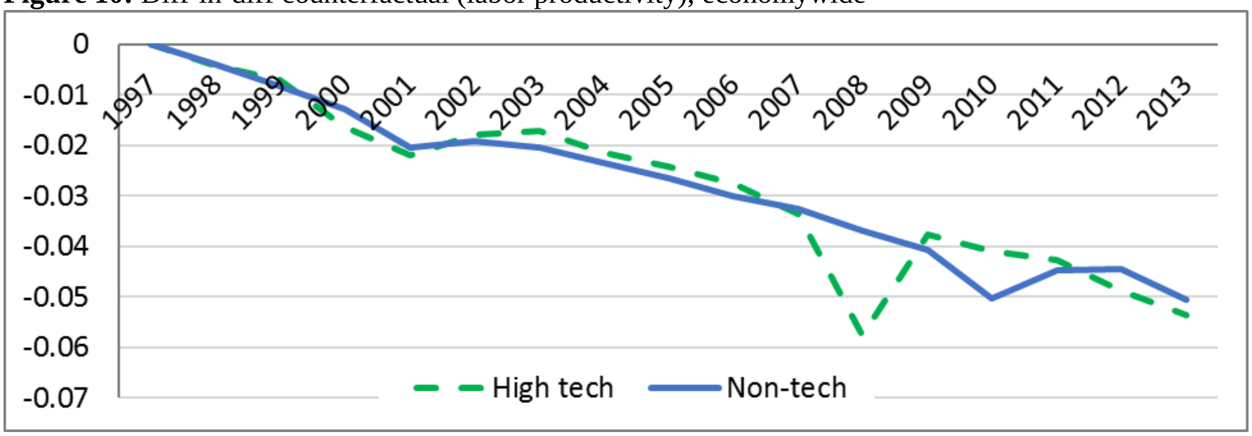
Note: Y axes do not begin at zero. Standard deviation of log labor productivity deviated from industry by year means. Young firms have age less than five. High tech is defined as in Hecker (2005). Author calculations from the RE-LBD. Finance, Insurance and Real Estate (NAICS 52-53) omitted.

Figure 9: Relative employment growth rates, high-productivity vs. average-productivity firm



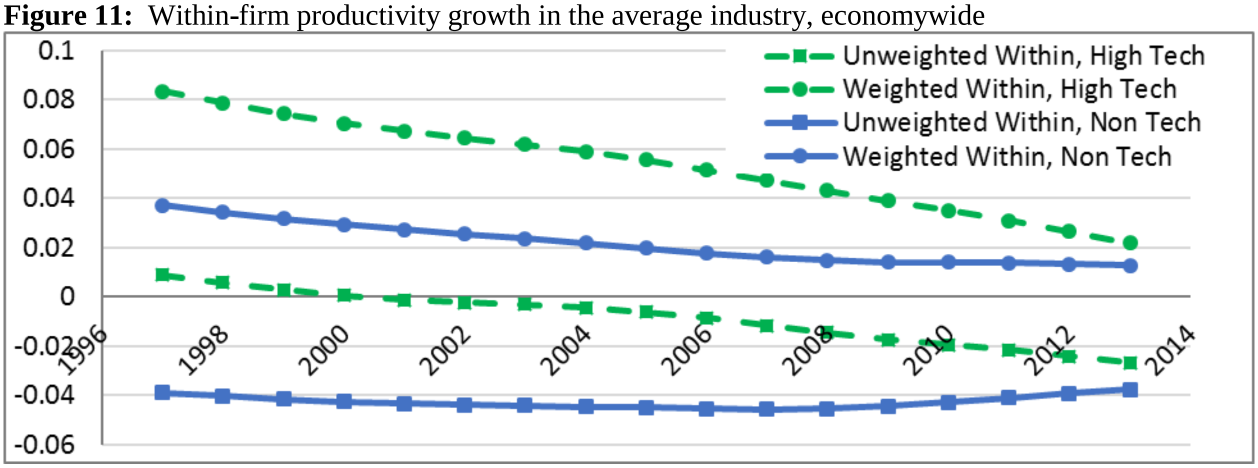
Note: Y axis does not start at zero. Growth rate of firm with labor productivity one std. dev. above industry mean vs. industry mean. Young firms have age less than five. High tech defined as in Hecker (2005). Author calculations from the RE-LBD. Finance, Insurance and Real Estate (NAICS 52-53) omitted.

Figure 10: Diff-in-diff counterfactual (labor productivity), economywide



Note: Figure depicts diff-in-diff counterfactual as described in the text. High tech is defined as in Hecker (2005). Author calculations from the RE-LBD.

Figure 11: Within-firm productivity growth in the average industry, economywide



Note: Average within-firm productivity growth, with and without employment weights. Author calculations from the RE-LBD.

Table 1: Effect of Lagged Productivity on Plant-Level Employment Growth and Exit

	Growth including exit		Exit	
	High tech	Non-tech	High tech	Non-tech
TFP*Young	0.2025***	0.2767***	-0.0292*	-0.0905***
	(0.0390)	(0.0090)	(0.0162)	(0.0037)
TFP*Young*Trend	0.0317***	0.0014	-0.0160***	-0.0005
	(0.0061)	(0.0014)	(0.0025)	(0.0006)
TFP*Young*Trend^2	-0.0012***	-0.00024***	0.0005***	0.0001***
	(0.0002)	(0.00005)	(0.0001)	(0.00002)
TFP*Mature	0.1228***	0.1439***	-0.0403***	-0.0464***
	(0.0174)	(0.0043)	(0.0072)	(0.0018)
TFP*Mature*Trend	0.0054**	0.0005	-0.0016	-0.0012
	(0.0026)	(0.0007)	(0.0011)	(0.0003)
TFP*Mature*Trend^2	-0.0003***	-0.00004*	0.0001***	0.00005***
	(0.0001)	(0.00002)	(0.00003)	(0.00001)
N	120000???	2000000???		
R^2				

Notes: Standard Errors in Parentheses. Dependent variable in Growth columns is DHS growth rate. Dependent variable in Exit columns is indicator=1 if exit, 0 Young firms have age less than 5. Unreported are estimates of controls including year effects, state effects, firm age dummies, firm size dummies, log plant level employment in period t, state cyclical indicators (change in state level unemployment rate), state cyclical indicators interacted with TFP. All variables that use TFP including all interactions are fully interacted with firm age dummies.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 2: Estimated Impact of Productivity on Plant-Level Equipment Investment Rate

	High tech	Non-tech
TFP*Young	0.0826***	-0.0125**
	(0.0236)	(0.0052)
TFP*Young*Trend	0.0189***	0.0156***
	(0.0037)	(0.0008)
TFP*Young*Trend^2	-0.0008***	-0.0004***
	(0.0001)	(0.0000)
TFP*Mature	0.0232**	0.0039
	(0.0105)	(0.0025)
TFP*Mature*Trend	0.0024	0.0067***
	(0.0016)	(0.0004)
TFP*Mature*Trend^2	-0.0001*	-0.0002***
	(0.00005)	(0.0000)
N		
R^2		

Notes: Standard errors in parentheses. Tech Sample is more than 120000 plant-year observations from 1981-2010. Young firms have age less than 5. Unreported are estimates of controls including year effects, state effects, firm age dummies, firm employment size dummies, log plant level employment in period t, dummies for initial capital, state cyclical indicators (change in state level unemployment rate), state cyclical indicators interacted with TFP. All variables that use TFP including all interactions are fully interacted with firm age dummies.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3: Lagged Labor Productivity and Firm-Level Employment Growth and Exit

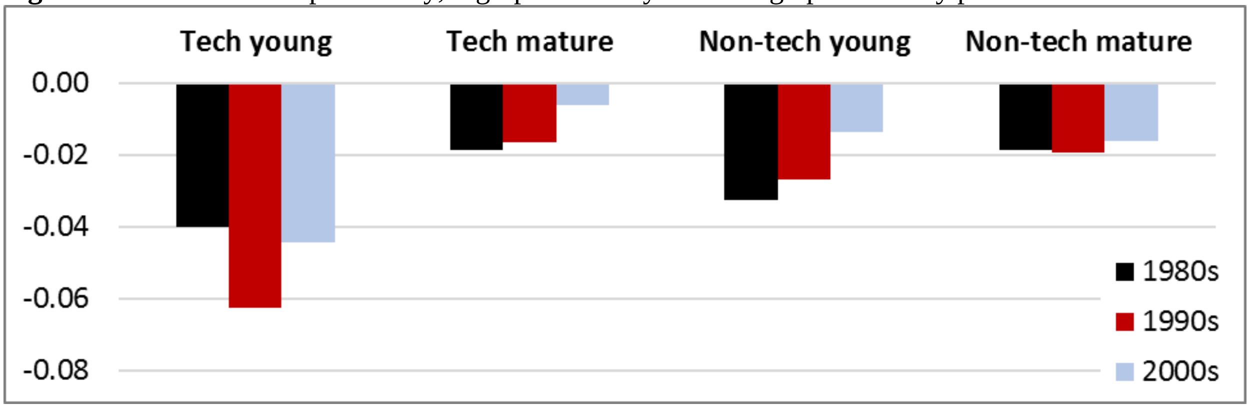
	Growth including exit		Exit	
	High tech	Non-tech	High tech	Non-tech
LP*Young	0.3845***	0.3467***	-0.1258***	-0.1224***
	0.0020	0.0005	0.0009	0.0002
LP*Young*Trend	-0.0141***	-0.0043***	0.0026***	0.0014***
	0.0006	0.0001	0.0002	0.0001
LP*Young*Trend^2	0.0004***	0.0000***	-0.0001***	0.0000***
	0.0000	0.0000	0.0000	0.0000
LP*Mature	0.2755***	0.2522***	-0.0710***	-0.0758***
	0.0021	0.0004	0.0009	0.0002
LP*Mature*Trend	-0.0042***	-0.0056***	-0.0008***	0.0020***
	0.0006	0.0001	0.0002	0.0000
LP*Mature*Trend^2	0.0000	0.0001***	0.0001***	-0.0001***
	0.0000	0.0000	0.0000	0.0000
N	55383000	55383000	55383000	55383000
R^2	0.1263	0.1083	0.1053	0.0931

Dependent variable in all regressions is firm-level employment growth rate (DHS). All regressions include controls for state business cycle (change in state unemployment rate) and firm employment size in period t-1. Labor productivity is measured as log difference from 6-digit NAICS industry mean. High tech is defined as in Hecker (2005). Observations rounded to nearest thousand.

*** p<0.01; ** p<0.05; * p<0.10

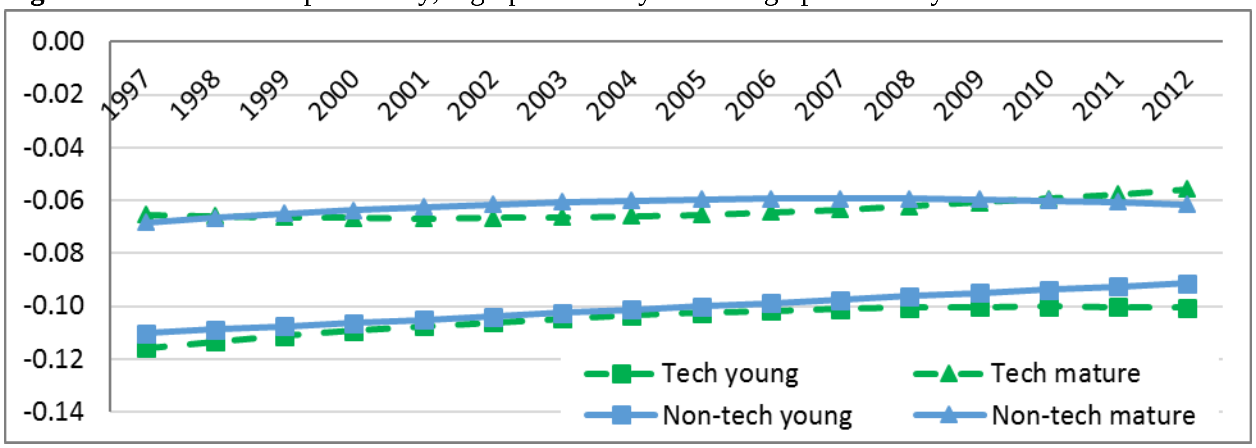
Appendix A. Figures and tables to supplement the main text

Figure A1: Relative exit probability, high-productivity vs. average-productivity plant



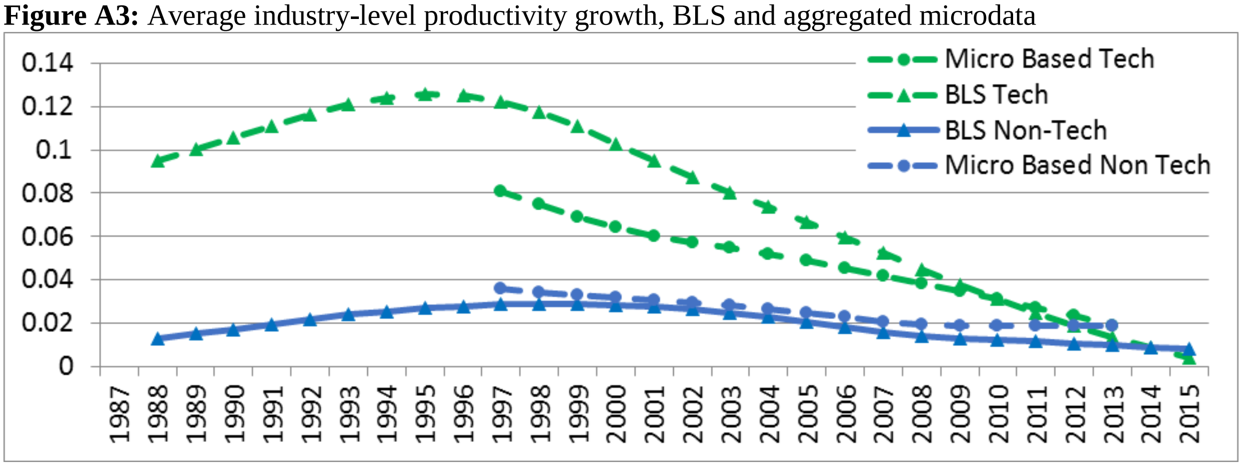
Note: Young firms have age less than 5. High tech is defined as in Hecker (2005). Exit probability of plant with TFP one std. dev. above industry mean vs. industry mean. Author calculations from the Longitudinal Business Database, the Annual Survey of Manufacturers, and the Census of Manufacturers.

Figure A2: Relative exit probability, high-productivity vs. average-productivity firm



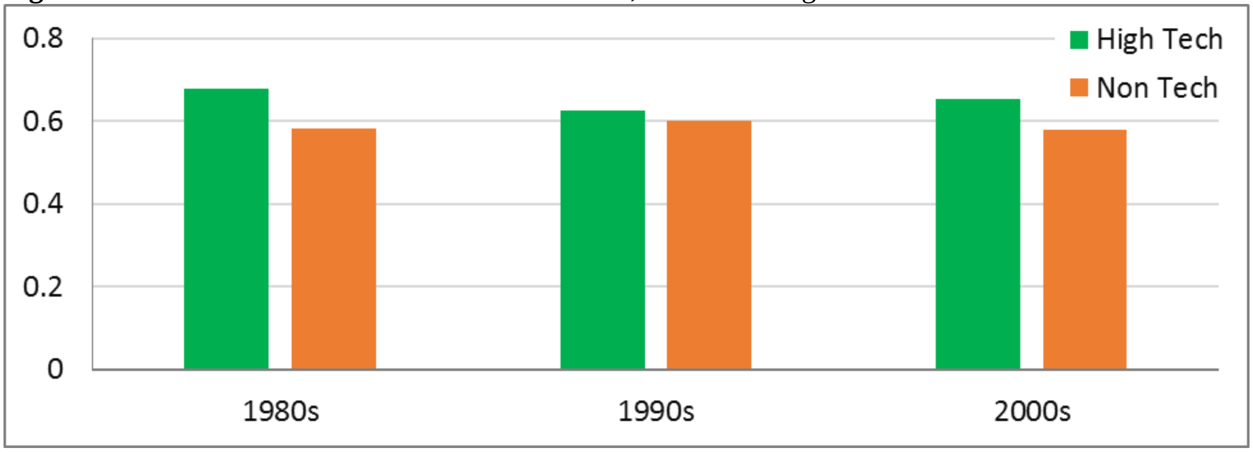
Note: Annual coefficients constructed from Table 3. Young firms have age less than five. High tech defined as in Hecker (2005). Exit probability of plant with labor productivity one std. dev. above industry mean vs. industry mean. Author calculations from the RE-LBD. Finance, Insurance and Real Estate (NAICS 52-53) omitted.

Figure A3: Average industry-level productivity growth, BLS and aggregated microdata



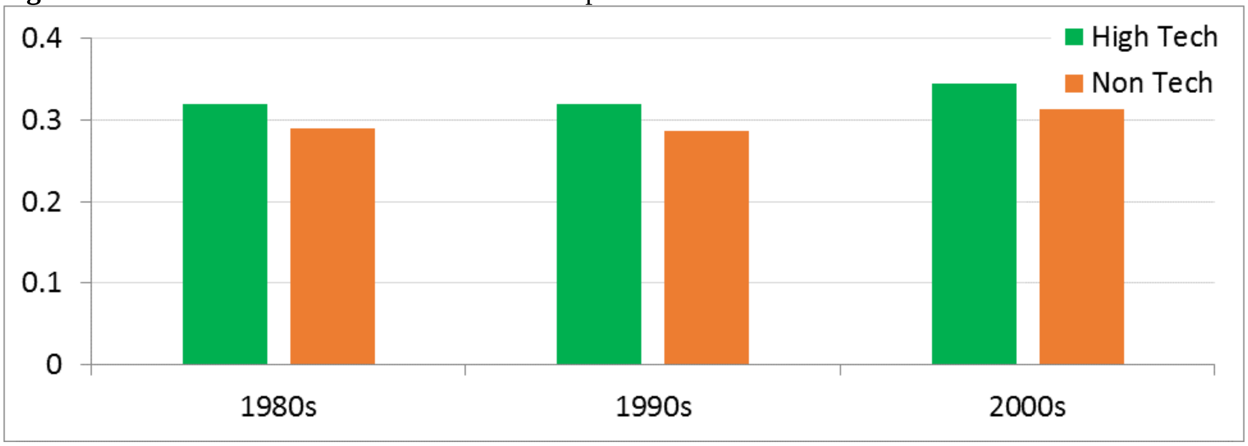
Source: BLS and author calculations from RE-LBD.

Figure A4: Persistence of establishment-level TFPR, manufacturing



Note: High tech is defined as in Hecker (2005). AR(1) coefficients for establishment TFPR, averaged by decade. The LBD-ASM-CM database is not ideally suited for estimating persistence since this requires relying on the longitudinal nature of the ASM/CM, which is less robust than the longitudinal properties of the LBD. That is, estimating productivity persistence parameters requires pairwise continuing plants in t and t+1 to be measured in the ASM/CM. The panel rotation of the ASM as well as Census years make this a challenge. That is, in the first years of a new ASM panel and in Census years we have a much smaller and less representative set of continuing plants than other years. For this exercise we exclude those years; even for other years, though, our propensity score weights are not ideally suited for making the sample of continuers representative. In principle, we can develop separate propensity score weights for this restricted sample of continuing plants. Doing so is more of a challenge, given the rotating nature of the ASM sample. See Figure B2 in appendix B for the same exercise on RPR productivity.

Figure A5: Standard deviation of innovations to plant-level TFPR



Note: High tech is defined as in Hecker (2005). For the set of years where we can estimate the AR(1) process (see note for Figure A4), we can also recover the distribution of innovations to plant-level TFP for continuing plants. Since this is for selected years we report averages of standard deviation of innovations to TFP by decade as we did with persistence.

Appendix B. Illustrative Model of Adjustment Costs

A. Model environment

Consider the following model of firm-level adjustment costs.[47] A firm maximizes the present discounted value of profits. The firm's value function and its components are specified as follows:

$$V(E_{et-1};A_{et}) = A_{et}E_{et}^{\varphi} - w_t E_{et} - C(H_{et}) + \beta V(E_{et+1};A_{et+1})$$

with:

$$C(H_{et}) = \begin{cases} \frac{\gamma}{2}\left(\frac{H_{et}}{E_{et}}\right)^2 & + F_{+}\max(H_{et},0) + F_{-}\max(-H_{et},0)\ if\ H_{et} \neq 0 \\ 0,\ otherwise & \end{cases}$$

$$a_{et} = \rho a_{et} + \eta_{et}$$
$$E_{et} = E_{et-1} + H_{et}$$

where $\varphi \leq 1$ due to product differentiation so that $A_{et}E_{et}^{\varphi}$ is the revenue function, E_{et} is employment for time t, H_{et} is net hires made at the beginning of time t, or $H_{et} = E_{et} - E_{e,t-1}$ (this can be positive or negative), w_t is the wage, and $a_{et} = \log(A_{et})$ is a revenue shock potentially reflecting TFPQ and demand shocks (for expositional convenience we focus on TFPQ). We interpret the revenue function curvature as reflecting product differentiation rather than decreasing returns to help draw out relations between revenue productivity and technical efficiency. That is, let firm-level prices be given by $P_{et} = Q_{et}^{\varphi-1}$ where $Q_{et} = \tilde{A}_{et}E_{et}$ is firm-level output subject to a CRTS technology. This implies that $A_{et} = \tilde{A}_{et}^{\varphi}$. In terms of the terminology of the literature and the main text, $\tilde{A}_{et}$ is TFPQ; and since labor is the only factor of production, both TFPR and revenue labor productivity (RLP) are given by $P_{et}\tilde{A}_{et}$. Note that this specification nests the price-taking version of the model with $\varphi = 1$. In that case, TFPQ, TFPR and RLP are equivalent We focus on the $\varphi < 1$ case in our calibration but discuss some aspects of the price-taking case below.

[47] We use the term "firm" for expositional purposes; for modeling purposes we do not distinguish between firms and establishments. Our main empirical results focus on establishments.

This simple adjustment cost model is similar to Cooper, Haltiwanger and Willis (2007, 2016) and Elsby and Michaels (2013) and, in principle, accommodates both convex and non-convex costs. Under non-convex costs the solution has the following form:

$$V = \max(V^I, V^H)$$

where

$$V^I = A_{et}E_{et-1}^{\varphi} - w_t E_{et-1} + \beta V(E_{et}; A_{et+1}) \text{ } if \text{ } H_{et} = 0$$

$$V^H = A_{et}E_{et}^{\varphi} - w_t E_{et} - C(H_{et}) + \beta V(E_{et}; A_{et+1}) \text{ } if \text{ } H_{et} \neq 0$$

with the notation indicating that V^I is the value of inaction (i.e., zero net hiring), and V^H is the value of nonzero net hiring (in either positive or negative amounts).

We view this model as illustrative and make no attempt to rigorously estimate the structural parameters since the model is missing some key features of the data. For example, we do not model entry or exit, and we do not have any lifecycle learning dynamics or frictions that make young firms different from more mature firms. Given these limitations, we regard the calibration as mostly providing guidance about the qualitative predictions for the key data moments we study.

Our main calibration exercise implements "general equilibrium" in the sense that the wage adjusts to clear the labor market; however, we fix the labor supply, so this calibration may be thought of as an extreme scenario. In unreported exercises, we also consider the opposite extreme in which labor supply is perfectly elastic and the wage is fixed (i.e., partial equilibrium). A limitation of the partial equilibrium exercise is that when the wage is fixed, adjustment frictions can have large effects on average firm size and therefore productivity via channels that are unrelated to reallocation. However, our key results on how adjustment costs affect reallocation rates, firm-level productivity responsiveness, and the OP covariance do not substantively depend on general vs. partial equilibrium. We report the more realistic (yet still extreme) general equilibrium (inelastic labor supply) results here.

B. Calibration

We set $\beta = 0.96$, consistent with annual data. We specify that $\varphi = 0.8$, consistent with a markup of 25 percent. For the shock process, we specify $\sigma_a = 0.35$, which is roughly the standard deviation of TFPR or RPR in U.S. manufacturing during the 1980s; and we set $\rho =$

0.65, broadly consistent with the AR(1) coefficient on TFPR and RPR that we find among manufacturing establishments in the 1980s (see Figure A4 and Appendix C). These values of σ_a and ρ imply that innovations to TFP have a standard deviation of $\sigma_\eta = 0.26$. Strictly speaking, if plant-level prices are endogenous (which this model permits) the appropriate empirical moments are those from RPR.

We calibrate the adjustment cost parameter(s) to target a job reallocation rate of 25 percent, roughly the rate for the U.S. manufacturing sector in the 1980s. Focusing only on kinked (non-convex) adjustment costs (i.e., setting $\gamma = 0$), we find that the target reallocation rate implies $F_+ = 0.85$ when $F_- = 0$. For quadratic adjustment costs (i.e., $F_+ = F_- = 0$) matching the job reallocation rate of 0.25 requires $\gamma = 1.3$. We make no attempt to jointly calibrate convex and non-convex adjustment costs for our illustrative purposes. For the key moments we study empirically, the model produces broadly similar predictions regardless of cost type. The literature suggests that non-convex costs are important for certain properties of microdata, so we focus on this cost type here and leave non-convex cost exercises unreported.

The model produces a (non-targeted) correlation between TFPQ and TFPR (which is the same as RLP in this one-factor setting) of 0.90, qualitatively similar to the 0.75 found by Foster, Haltiwanger and Syverson. This strong correlation implies that the responsiveness of growth to realizations of productivity is essentially the same whether we use TFPQ or TFPR/RLP as the measure of productivity.

We consider two types of experiments in the simulation. The first is an increase in adjustment frictions from $F_+ = 0.85, F_- = 0$, our baseline non-convex cost calibration described above. We then increase F_- from zero to study rising adjustment costs from a starting point that matches the patterns of TFP and reallocation in the 1980s. In the second experiment, we fix adjustment costs at the baseline but vary TFP dispersion around the baseline to study the relation between TFP dispersion and responsiveness, RLP dispersion, and the OP covariance.

C. Changing adjustment costs

Here we describe results from the general equilibrium (i.e., flexible wage with inelastic labor supply) model. We consider several moments calculated on model-simulated data;[48] these moments are the job reallocation rate; the standard deviation of labor productivity (or,

[48] We simulate 2000 firms for 1000 periods then discard the first 100 periods.

equivalently in this model, TFPR); the TFP coefficient from a regression of firm-level employment growth from time t to $t+1$ on TFP in time t and (log) employment in time t (that is, we run a regression analogous to equation (4) in the paper; hereafter we call this moment the "productivity coefficient"); the Olley-Pakes covariance for TFP (hereafter "OP covariance"); and the OP covariance for labor productivity.

Figures B1 and B2 show the effect of increasing adjustment frictions (i.e., raising the downsizing cost F_{-} while holding fixed $F_{+} = 0.85$). An increase in adjustment frictions yields: (i) a decline in the job reallocation rate; (ii) a decline in the productivity coefficient; (iii) an increase in the dispersion of RLP; and (iv) a decline in the OP covariance for both TFP and RLP (where employment serves as weights). Each of these relationships is monotonic in adjustment costs except the Olley-Pakes covariance for RLP (which we discuss below).

There are many possible moments relating growth to TFP that are similarly sensitive to adjustment costs. For example, in unreported regression results (in which we always control for period-t log employment), increasing adjustment costs yield (i) a decline in the estimated coefficient of a regression of firm-level growth between t and $t+1$ on TFP in period $t+1$ (rather than t); (ii) a decline in the estimated coefficient of a regression of firm-level growth between t and $t+1$ on the change in TFP from period t to $t+1$; and (iii) a decline in the estimated coefficient of a regression of firm-level growth between t and $t+1$ on the innovation of TFP (η_{et}). In principle, we could use any of these moments to detect a change in adjustment frictions. We use the specification reported in Figure B1 for measurement and econometric reasons discussed in the main text. The exact timing in the model vs. the data are different, so it is reassuring that the predictions on responsiveness hold equally well qualitatively in the numerical analysis using current or lagged productivity.

The model-based predictions in B1 and B2 are the primary moments that we explore empirically in the main text. In the empirical analysis we also consider the estimated coefficient of firm-level growth between t and $t+1$ on revenue labor productivity in t (with log period-t employment as a control as usual). Given the simple revenue functions in the model, this estimated coefficient is identical to the coefficient on TFP shown on Figure B1 (this is because the only production factor is labor, and employment is explicitly included in all regressions). This precise equivalence is model dependent, but the general inference is not. That is, in response to an increase in adjustment frictions, there should be a decline in the covariance between firm-level growth and realizations of labor productivity as firms find it more costly to equalize their marginal products.

D. Changing TFP dispersion

Figure B3 shows how key moments vary with TFP dispersion. Increased TFP dispersion yields: (i) an increase in the job reallocation rate; (ii) an increase in the productivity coefficient; and (iii) an increase in the dispersion of labor productivity. As before, the finding for the productivity coefficient also holds using real labor productivity as the regressor.

While the results in Figure B3 are generally intuitive, one finding merits further discussion—specifically, the finding that responsiveness increases in TFP dispersion. The net effect of TFP dispersion on responsiveness reflects two competing mechanisms (as discussed in the main text). The first is the "real options" effect: Non-convex costs create "inaction bands" or regions of the TFP innovation range in which firms prefer inaction (i.e., zero hiring) to action. Inaction bands widen as shock dispersion or volatility rises (consistent with an "uncertainty" interpretation), which, *ceteris paribus*, reduces responsiveness. We observe this effect in our simulated data when we examine only the extensive margin: a given absolute change in TFP is more likely to induce action when TFP dispersion is smaller (holding initial employment constant). However, in the model this effect is dominated by the "volatility effect" in which adjustments—when they actually do occur—are larger when TFP is more widely dispersed.

E. The frictionless case

The patterns in Figures B1-B3 are, for the most part, robust to changes in the curvature of the revenue function, the shock space and the adjustment cost parameters. One exception highlights the importance of using multiple moments in our empirical exercises. Specifically, in a frictionless benchmark with zero adjustment costs (in contrast to our baseline above, in which hiring costs are set to $F_{+} = 0.85$), there is zero labor productivity dispersion and, therefore, zero OP covariance for labor productivity (though still positive OP covariance for TFP). At first glance, this implies that the OP covariance for labor productivity may not be an informative moment, but we show here that the frictionless benchmark yields patterns that are very far from empirical plausibility.

Figure B4 shows the effects of increasing adjustment costs from zero, the frictionless case. As adjustment frictions rise above zero, labor productivity dispersion rises (Figure B4) and, consistent with the discussion above, reallocation and the productivity coefficient decline.

But Figure B5 shows that the OP covariance for labor productivity initially rises as labor productivity begins to be dispersed, continuing to rise over the range of adjustment frictions that produce reallocation rates above 30 percent (compare Figure B5 to Figure B4). But since productivity responsiveness declines monotonically as adjustment costs rise, the OP covariance eventually declines as labor is increasingly "trapped" in unproductive firms while productive firms are starved of resources (i.e., employment weight). Thus, the OP covariance for labor productivity is decreasing in adjustment costs (and increasing in misallocation) across the plausible range of costs. This pattern is related to that found in Bartelsman, Haltiwanger and Scarpetta (2013), in whose model distortions reduce the OP covariance for labor productivity as long as the benchmark is characterized by sufficient frictions.

As can be seen on Figure B4, the frictionless benchmark produces implausibly large rates of reallocation—above 100 percent, while empirical reallocation in the manufacturing sector was around 25 percent in the 1980s and has since declined. Here we demonstrate basic analytical intuition underlying the empirical implausibility of the frictionless case. In the frictionless case ($\gamma = F_{+} = F_{-} = 0$), the first-order condition for labor is given by:

$$E_{et} = \left(\frac{\varphi A_{et}}{w_t}\right)^{\frac{1}{1-\varphi}}$$

Taking logs (indicated by lower case) and differences (indicated by Δ) and sweeping out year and industry effects yields:

$$\Delta e_{et} = \frac{1}{1-\varphi} \Delta a_{et}$$

which implies

$$std(\Delta e_{et}) = \frac{1}{1-\varphi} std(\Delta a_{et})$$

where $std()$ indicates standard deviation. That is, in the frictionless model, the dispersion of employment growth rates (i.e., log differences) is proportional to the dispersion of TFP with the factor of proportionality greater than one. The $std(\Delta a_{et})$ among continuing manufacturing plants is about 0.33 (this is from RPR – similar statistics emerge from TFPR). For $\varphi = 0.8$, corresponding to a markup of 25 percent, we should expect $std(\Delta e_{et}) = 1.65$; for a 33 percent markup we should expect $std(\Delta e_{et}) = 1.32$. Yet in U.S. manufacturing data $std(\Delta e_{et}) = 0.35$.

This relatively low dispersion of plant-level employment growth rates compared to the dispersion in shocks illustrates that the frictionless model yields implausible empirical patterns.

Given the empirical implausibility of the frictionless model, then, we are comfortable drawing inference from key moments—including the OP covariance for labor productivity—along the plausible range of adjustment costs. We elaborate further on strengths and limitations of the OP decomposition below.

F. *The Olley-Pakes decomposition and covariance*

While we show above that the OP covariance is declining in adjustment costs over plausible parameterizations, additional doubts about its usefulness may arise from limitations of the accounting-based decomposition on which it is based (Petrin and Levinsohn (2012); Petrin, White and Reiter (2011); Hsieh and Klenow (2017)). We do not formally use the OP covariance in any of our empirical analysis but instead use diff-in-diff counterfactuals. However, it is instructive to compare and contrast our diff-in-diff counterfactuals with the OP covariance. As noted in the text, the OP productivity decomposition is given by:[49]

$$A_t^{OP} = \sum_f \theta_{et} A_{et} = \bar{A}_t + cov(\theta_{et}, A_{et})$$

where A_t^{OP} is the Olley-Pakes concept of industry aggregate productivity, θ_{et} is the employment share of firm f, A_{et} is productivity for firm f, $\bar{A}_t$ is (unweighted) average productivity for the industry, and $cov(\theta_{et}, A_{et})$ is the OP covariance (which is proportional to the technical definition of covariance).

Define aggregate productivity as $A_t = Q_t/E_t$, that is, aggregate output divided by aggregate employment. A critical question is how well A_t^{OP} matches A_t; a related question is how to interpret the quantitative variation in the OP covariance term. Continuing with our model, characterized by constant returns to scale but potentially imperfect competition,[50] note that aggregate output is given by:

[49] In empirical work, it is common to use a weighted average of log firm-level productivity to avoid index number problems. We use the weighted average of firm-level productivity here to highlight that under CRTS and price-taking behavior that this weighted average is equal to aggregate productivity. Empirical implementation still requires addressing the index number problems.

[50] A broader interpretation of our simple model is that the revenue function curvature reflects either (or both) imperfect competition or decreasing returns to scale. The effects of either on the accuracy of the OP decomposition are similar, so we proceed in this subsection with a focus on the imperfect competition interpretation while noting that the decreasing returns interpretation has similar implications.

$$Q_t = \left(\sum_f Q_{ft}^{\varphi}\right)^{\frac{1}{\varphi}}$$

where $\varphi \leq 1$. Aggregate productivity is as follows:

$$A_t = \frac{Q_t}{E_t} = \frac{\left(\sum_f A_{et} E_{et}^{\varphi}\right)^{\frac{1}{\varphi}}}{E_t}$$

It is straightforward to see that if $\varphi = 1$, that is, perfect competition, aggregate output of the final good is simply the sum of firm-level output, and the OP decomposition measures aggregate productivity exactly:

$$A_t = \frac{Q_t}{E_t} = \frac{\sum_f A_{ft} E_{ft}}{E_t} = \sum_f A_{ft} \theta_{ft} = A_t^{OP}$$

In this case the size distribution can only be determined in the presence of frictions; but the OP covariance term's effect on aggregate productivity is intuitive since the firm with the highest average productivity also has the highest marginal productivity. Moving resources to the most productive firm will always increase aggregate productivity (this inference is robust to a model with multiple inputs). The key is that with CRTS and perfect competition, the marginal revenue product of a firm does not change with scale but only varies with TFPQ.

For $\varphi < 1$, however, aggregate productivity is no longer equal to the OP weighted average of firm productivity. Moreover, it is no longer the case that continually moving resources to the most-productive firm (in terms of average productivity) will increase aggregate productivity. This is because with $\varphi < 1$, as resources are moved to the most productive firms marginal revenue products rise for the least productive firms and fall for the most productive firms. This implies that with $\varphi < 1$ the OP weighted average of firm productivity declines more rapidly with an increase in adjustment costs than aggregate productivity.

To overcome this limitation of the OP covariance measure, in our empirical analysis we rely instead on a diff-in-diff counterfactual (equation (5) in the text) given by:

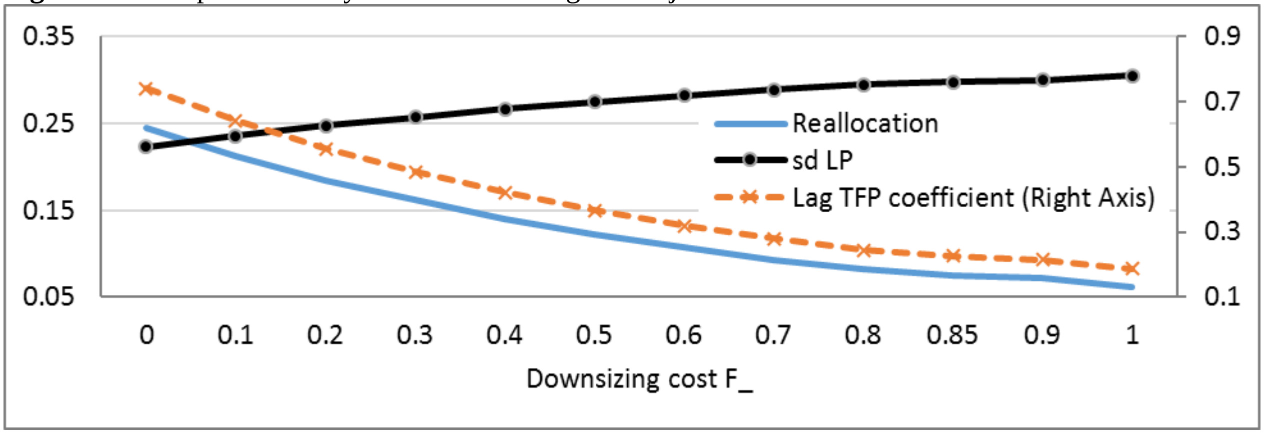
$$\Delta_t^{t+1} = \sum_f (\theta_{e,t+1}^{T} - \theta_{e,t+1}^{NT}) a_{et}$$

This diff-in-diff counterfactual isolates the impact of changing responsiveness on the weighted mean holding everything else constant.

Figure B6 illustrates how aggregate productivity, the OP covariance using TFPQ and the diff-in-diff counterfactuals using either TFPQ or RLP (equivalent to TFPR in the one-factor model). Consistent with the discussion above, aggregate productivity declines less rapidly than the OP covariance with an increase in adjustment costs. In contrast, the diff-in-diff counterfactuals track aggregate productivity very closely. In unreported results, we have shown that the properties of Figure B6 are robust to alternative values of φ. In particular, the close quantitative correspondence between changes in aggregate productivity and the diff-in-diff counterfactuals is robust to alternative values of $\varphi < 1$.

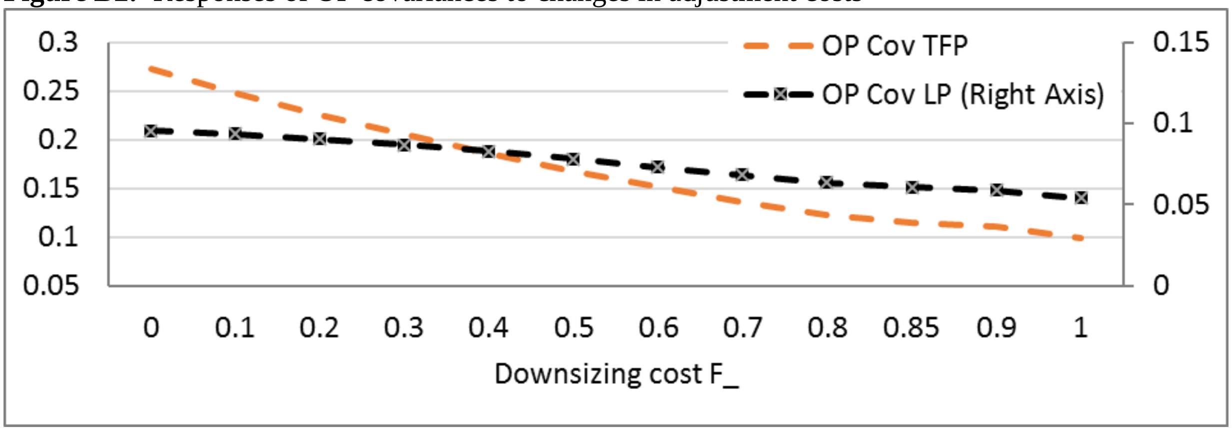
In closing, one might ask why we don't just quantify the changes in aggregate (industry) productivity in our empirical analysis. The reason is this would not isolate the changes due to changing adjustment frictions. Our diff-in-diff counterfactual achieves this objective.

Figure B1: Responses of key moments to changes in adjustment costs



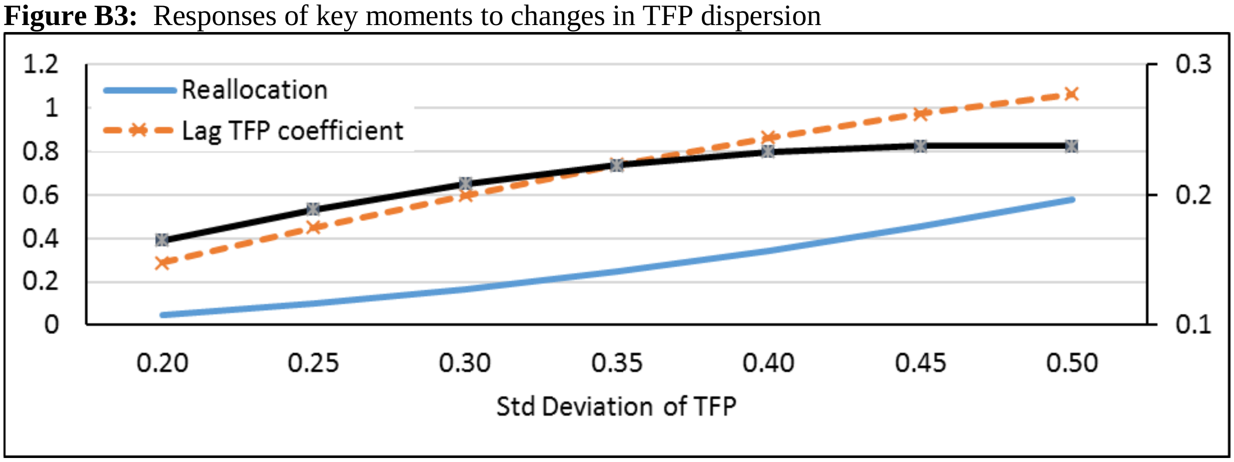
Note: The x axis reflects values of F_{-}, or the cost of reducing employment, holding the hiring cost F_{+} fixed at $F_{+} = 0.85$. General equilibrium model with flexible wage and inelastic labor supply.

Figure B2: Responses of OP covariances to changes in adjustment costs



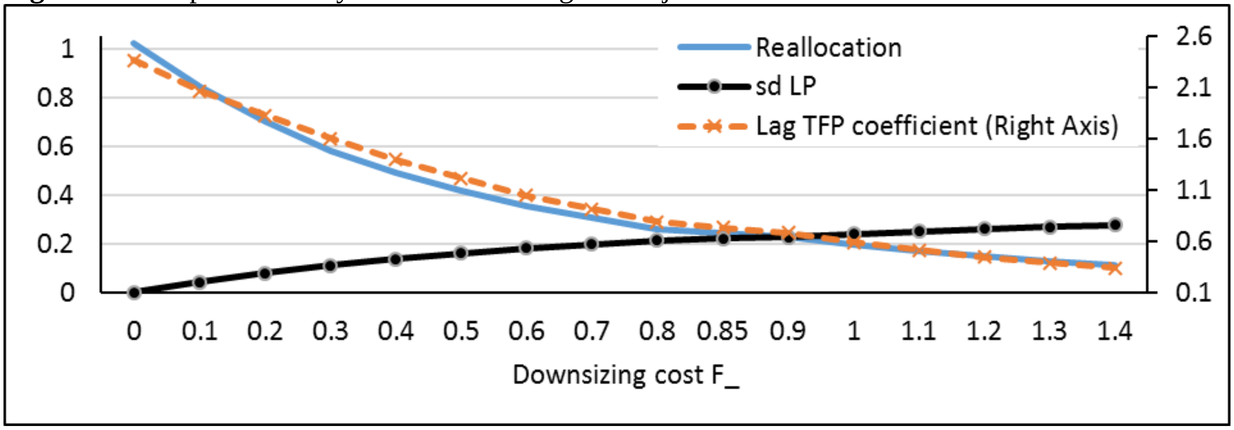
Note: The x axis reflects values of F_{-}, or the cost of reducing employment, holding the hiring cost F_{+} fixed at $F_{+} = 0.85$. General equilibrium model with flexible wage and inelastic labor supply.

Figure B3: Responses of key moments to changes in TFP dispersion



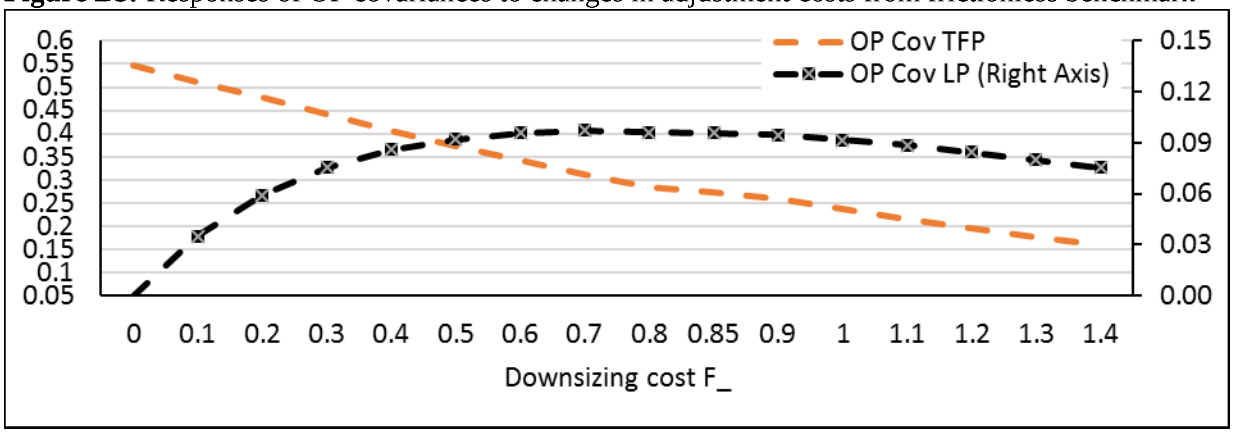
Note: Model with kinked adjustment costs ($F_{+} = 0.85, F_{-} = 0$). General equilibrium model with flexible wage and inelastic labor supply.

Figure B4: Responses of key moments to changes in adjustment costs from frictionless benchmark



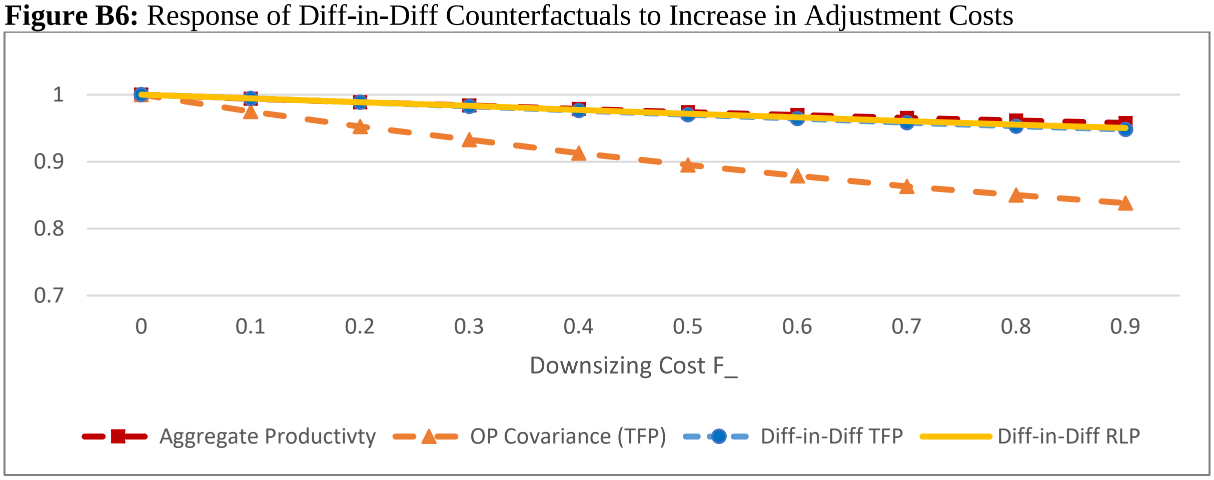
Note: Model with no upward or downward adjustment costs ($F_+ = 0$) with varying downward adjustment costs (F_-) indicated on the x axis. General equilibrium model with flexible wage and inelastic labor supply.

Figure B5: Responses of OP covariances to changes in adjustment costs from frictionless benchmark



Note: Model with no upward adjustment costs ($F_+ = 0$) with varying downward adjustment costs (F_-) indicated on the x axis. General equilibrium model with flexible wage and inelastic labor supply.

Figure B6: Response of Diff-in-Diff Counterfactuals to Increase in Adjustment Costs



Note: The x axis reflects values of F_-, or the cost of reducing employment, holding the hiring cost F_+ fixed at $F_+ = 0.85$. General equilibrium model with flexible wage and inelastic labor supply.

Appendix C. Alternative TFP calculation

While our TFPR measure as a measure of TFP is common in related literature, as we discus in the main text we also consider an estimate of RPR using the proxy method of Wooldridge (2009). As we show in equation (3), RPR is only a function of exogenous TFPQ and demand shocks (even if plant-level prices are endogenous) because the elasticities recovered by revenue function estimation are revenue elasticities (not factor elasticities) capturing both production and demand parameters (Foster et al. (2017)). In this appendix, we discuss the estimation of RPR and the results using the RPR measure of TFP. Given the possible presence of demand shocks, RPR should be interpreted as reflecting both TFPQ and demand shocks.

Foster et al. (2017) find that the Woolridge residuals are sensitive to outliers; pooling across a large number of observations mitigates this sensitivity, so we estimate revenue elasticities that vary at the 3-digit NAICS level.[51] After estimating the elasticities, we compute the revenue productivity residuals and deviate the latter from 6-digit NAICS industry by year means. We find that RPR_{et} has a correlation of 0.76 with $TFPR_{et}$.

We replicate our main empirical exercises replacing TFPR with RPR. Figure C1 shows the evolution of within-industry dispersion in RPR_{et} for manufacturing plants. Consistent with Figure 4, we observe gradually rising RPR dispersion throughout the time period, with higher dispersion in high tech than elsewhere. Figure C2 reports AR(1) coefficients for plant-level RPR (see note to Figure A4 in appendix A for a discussion of this measure and its limitations in our dataset). Again, RPR results are consistent with TFPR results, confirming that changes in the TFP distribution cannot explain aggregate job reallocation patterns.

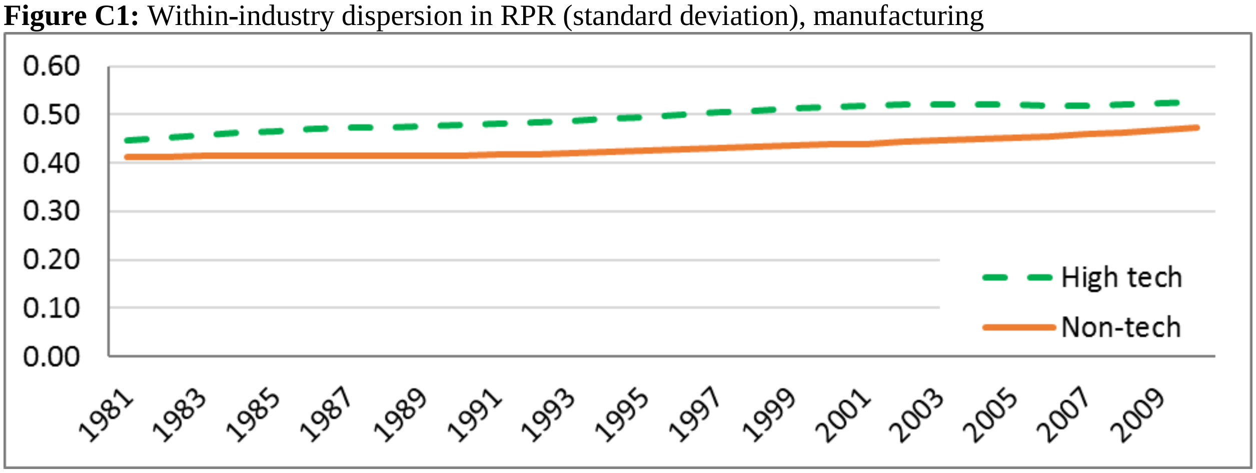
We estimate equation (4) using RPR in place of TFP. Figure C3 reports growth differentials (between the plant with productivity one standard deviation above its industry mean and the mean) as discussed in the main text. The results are generally consistent with those reported for TFP on Figure 5, with young firm productivity responsiveness in high tech that rises from the 1980s to the 1990s then falls in the 2000s. Among mature firms in high tech,

[51] The cost share-based TFPR measure we use in the main text is constructed with cost shares that vary at the 4-digit SIC level prior to 1997 and the 6-digit NAICS level thereafter. The instability and outlier sensitivity of the RPR elasticity estimates precludes this level of detail. The cost share-based TFPR method therefore allows for more flexibility in elasticity values implying a better fit in detailed industries, while the RPR method avoids problems of price endogeneity and isolates exogenous TFPQ and demand shocks. This tradeoff is the reason we ensure robustness of our exercises to both productivity concepts.

responsiveness is somewhat flat from the 1980s to the 1990s before falling markedly in the 2000s.

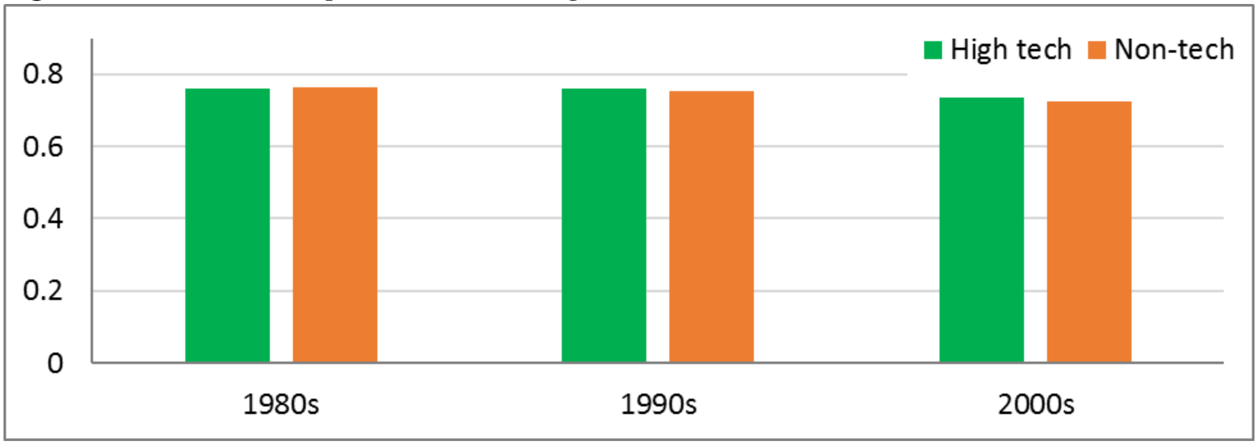
Finally, Figure C4 reports the diff-in-diff counterfactual described by equation (5). Among high tech plants, declining responsiveness produces a counterfactual that is broadly similar—both qualitatively and quantitatively—with the TFPR-based results from Figure 6, with a productivity "drag" that is only slightly smaller under RPR than under TFPR. Among non-tech plants, the counterfactual produces somewhat different results from those reported in Figure 6, with a gap opening up early in the sample then remaining stable (and negative) after the late 1990s. In general the RPR results confirm the TFP-based findings suggesting a quantitatively significant change in the contribution of reallocation to aggregate productivity growth.

Figure C1: Within-industry dispersion in RPR (standard deviation), manufacturing



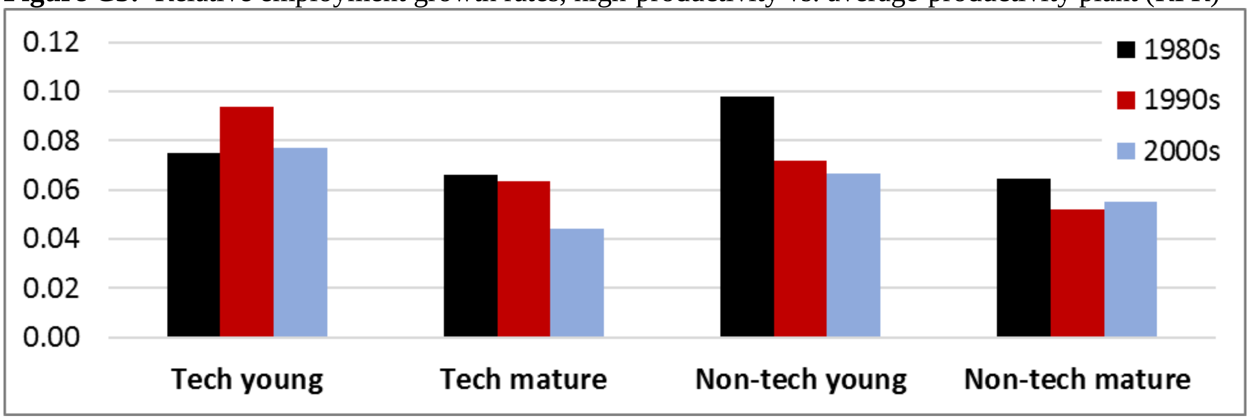
Note: The standard deviation is the based on within-detailed industry log revenue productivity residual. High tech is defined as in Hecker (2005). Manufacturing is defined on a consistent NAICS basis. Author calculations from the Longitudinal Business Database, the Annual Survey of Manufacturers, and the Census of Manufacturers. Hodrick Prescott Trends depicted.

Figure C2: Persistence of plant-level RPR: High tech vs. non-tech



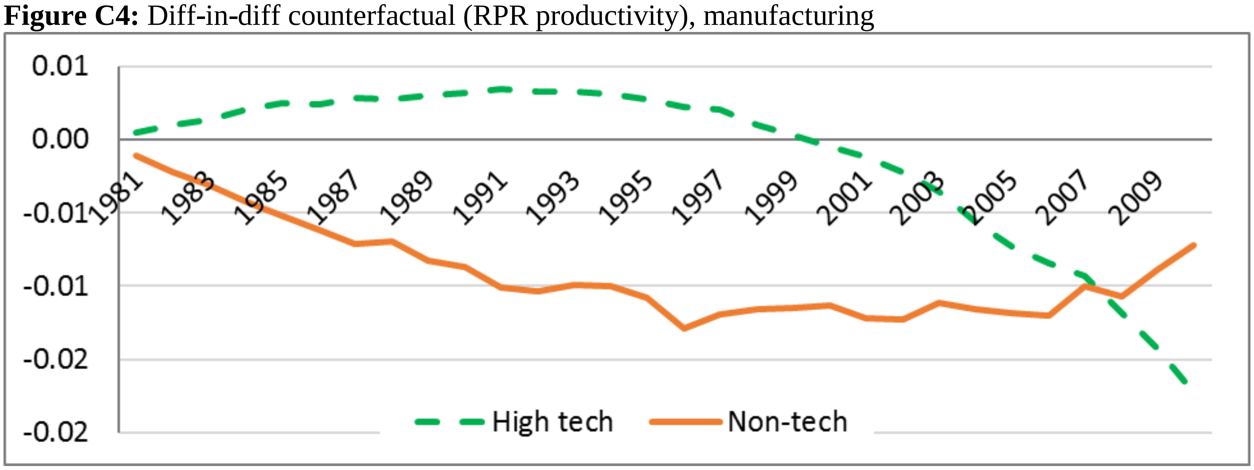
Note: High tech is defined as in Hecker (2005). Author calculations from the Longitudinal Business Database, the Annual Survey of Manufacturers, and the Census of Manufacturers.

Figure C3: Relative employment growth rates, high-productivity vs. average-productivity plant (RPR)



Note: Young firms have age less than 5. High tech is defined as in Hecker (2005). Author calculations from the Longitudinal Business Database, the Annual Survey of Manufacturers, and the Census of Manufacturers.

Figure C4: Diff-in-diff counterfactual (RPR productivity), manufacturing



Note: Figure depicts diff-in-diff counterfactual as described in the text. High tech is defined as in Hecker (2005). Author calculations from the Longitudinal Business Database, the Annual Survey of Manufacturers, and the Census of Manufacturers.

Appendix D. Changing Business Models in Manufacturing

A. Investment

As noted in the text, Table 2 and Figure 5 report our standard responsiveness regression using establishment investment rates as the dependent variable in place of employment growth. We include the capital stock as an additional state variable in these regressions; we do not include the capital stock in our main employment growth regressions, but in unreported exercises we find our main results are robust to its inclusion. The timing is slightly different for equipment investment as opposed to the employment growth specifications: in the employment growth specifications, we measure growth from March of t to March of t+1 as a function of size in March t and productivity for year t. In the investment specification, we measure investment throughout year t as a function of size in March t, productivity for year t, and capital stock at the beginning of year t. Consistent with standard models, we have in mind a time-to-build assumption that investment during period t contributes to capital used during period t+1, which means the difference in timing from the employment growth regressions is not large. Moreover, our model exercises in Appendix B suggest that our theoretical framework is not heavily dependent on specific timing concerns.

B. Globalization

Globalization may be playing a role in declining responsiveness since increased exposure to foreign trade facilitates adjustment by scaling international operations. That is, it may be that rather than growing domestically, productive firms are more likely to expand and produce in other countries, a dynamic that could eliminate or even reverse the standard positive correlation between growth and productivity (since we do not observe employment outside the U.S.). There is substantial evidence already that the decline in US manufacturing employment is closely linked to rising import penetration of production activity from low wage countries (see, e.g., Bernard, Jensen and Schott (2006), Schott (2008) and Pierce and Schott (2016)).

Bernard, Jensen and Schott (2006) and Schott (2006) develop measures of import penetration ratios from low wage countries. Their measures vary by 4-digit SIC industry from 1972-2005 and by 6-digit NAICS industry from 1989-2005; we extend the time series using the public domain information from Census on imports by country and industry.[52] We integrate

[52] To construct low-wage import penetration data by year and industry, Bernard, Jensen and Schott first construct domestic absorption for each industry. Next, they construct total imports of goods produced by each industry that are sourced in a low-wage country, which are defined as countries whose GDP per capita is less than 5 percent of the

these public domain data into our data infrastructure from 1981-2010. Our ability to integrate this is facilitated by our having 4-digit SIC codes in the micro level data from 1981-1996 and 6-digit NAICS codes from 1981-2010; hence, we need not rely on aggregate SIC/NAICS concordances.[53] Figure D1 shows aggregate import penetration ratios in and out of high tech manufacturing.

Table D1 presents results of a modified version of our main regressions in equation (3). The additional regressors added are the 6-digit NAICS import penetration ratio for each year and the interaction of this ratio with lagged TFP. We permit the coefficients on this interaction effect to differ between plants belonging to young and mature firms. The main effect of the import penetration (not reported) is negative and significant: Consistent with Bernard, Jensen and Schott (2006), plants in industries with especially large increases in import penetration have lower net employment growth.

The last two rows of Table D1 show that the interaction effect for young plants of lagged TFP and the import penetration ratio is estimated to be negative and significant. This implies that young-firm plants in industries with especially large increases in import penetration ratios have larger decreases in responsiveness. In Figure D2, we quantify the effect of changing import penetration ratios using the estimated effects from Table D1. The overall effects show, consistent with Table 1, that the marginal effect of productivity on employment growth among young high tech firms increased from the 1980s to 1990s then declined in the post-2000 period. We compute the fraction of these patterns accounted for by the changing import penetration ratios by using the coefficients from Table D1 along with the aggregate pattern of import penetration ratios for high tech manufacturing. The role of rising penetration is very modest in the 1980s to 1990s. However, the rapid rise in import penetration during the 2000s accounts for a substantial share (about 16 percent) of the overall decline in responsiveness over that period. We also perform these analyses using the Wooldridge (2009) RPR productivity measure (unreported), finding no significant role for import penetration, so we consider this evidence mixed. More research on globalization and dynamism is needed; promising avenues include

U.S. Import penetration is the ratio of low-wage imports to total domestic absorption, by industry and year. We thank Peter Schott for providing the import data and guidance necessary for extending the dataset.

[53] We integrate the SIC-based import penetration ratios from 1981-88 and the NAICS-based ratios from 1989-2010 into the micro data. We use the internally consistent NAICS codes in the micro data from 1981-2010 to conduct our analysis. (see Fort and Klimek (2016)).

specific policy variation, distinction between intermediate and final goods competition, and differences between TFP concepts.

C. Composition Effects

In the high tech manufacturing sector, another possible cause of declining productivity responsiveness after 2000 is the transition from "general-purpose" to "special-purpose" equipment manufacturing in the U.S documented by Byrne (2015).[54] Perhaps manufacturers of special-purpose products are less responsive to productivity shocks due to demand constraints or uncompetitive environments that reduce adjustment imperatives. Figure D3 shows that during the 1990s the share of employment in among general purpose technology producers grew rapidly but, consistent with Byrne (2015) (who examined revenue shares), the general purpose share has fallen substantially since the late 1990s. Given these compositional changes, it is possible that changing average responsiveness reflects differential responsiveness across industries.

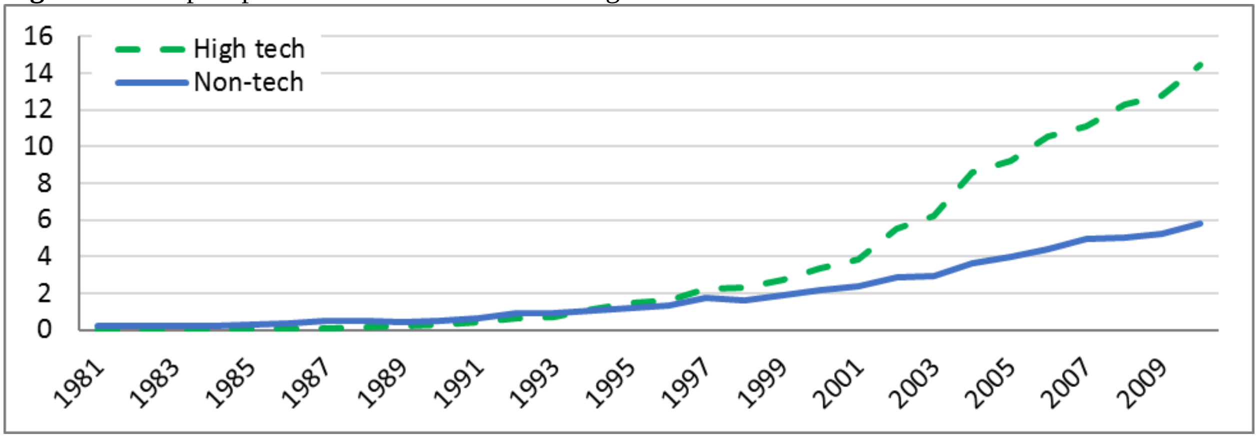
We estimate equation (4) separately for each 6-digit industry in high tech manufacturing but, importantly, we omit time trend interactions from the specification. With the estimated responsiveness coefficients for each 6-digit industry, we compute the employment-weighted aggregate responsiveness in each year using the actual annual 6-digit employment weights.[55] If any industry composition effect—including the shift between general and special purpose electronics—is driving our results, we should see these aggregated responsiveness patterns mimicking the result from Figure 5.

Figure D4 shows the implied changing responsiveness over time due to composition effects within high tech manufacturing. There is no implied increase in responsiveness due to composition effects from the 1980s to the 1990s (which would have been expected if general-purpose producers were more responsive on average), and there is actually a modest increase in responsiveness from the 1990s to the 2000s rather than a decline. Declining responsiveness must therefore be a within-category phenomenon with respect to the general-purpose/special-purpose taxonomy and other industry characteristics.

[54] We thank Christopher Foote for this insight.

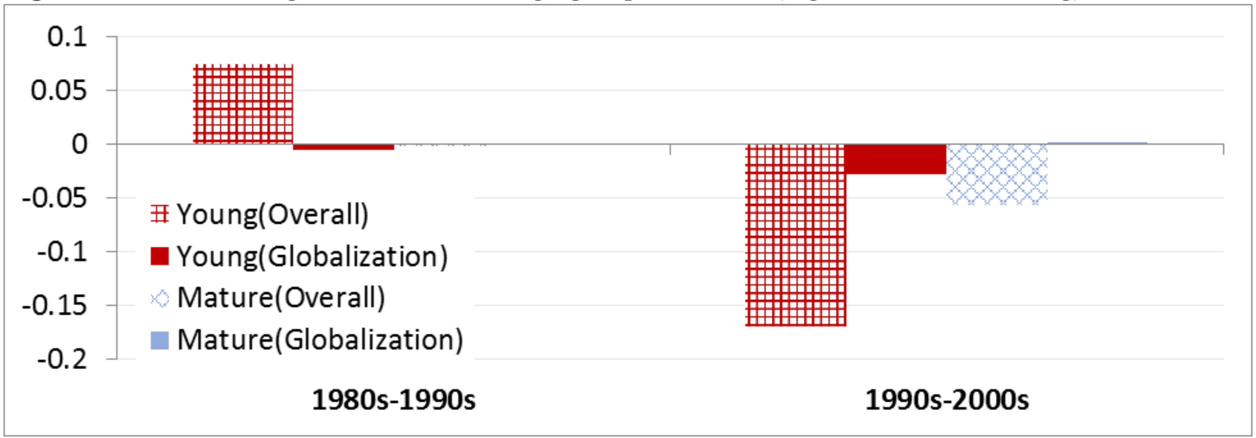
[55] We use employment weights given our interest in the implications of changing responsiveness for job reallocation.

Figure D1: Import penetration ratios from low-wage countries



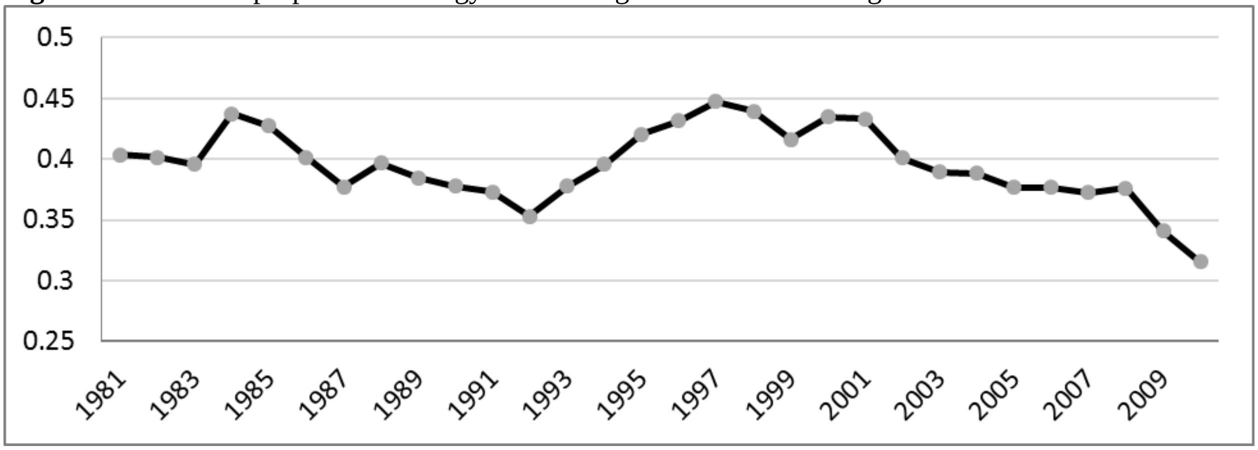
Source: Extended versions of Import Penetration Ratios from Bernard, Jensen and Schott (2006) and Schott (2008). Reported statistics are averages across 6-digit NAICS industries for high tech and Non tech industries.

Figure D2: The role of globalization in changing responsiveness (high tech Manufacturing)



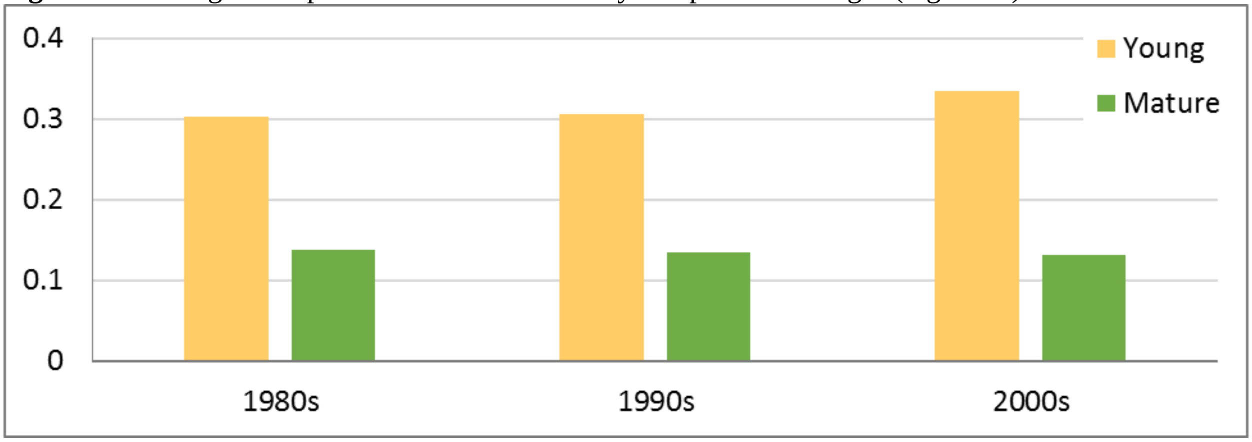
Note: "Overall" bars for young and mature are the change in marginal responsiveness of employment growth to productivity across decades. Globalization reflects implied change in marginal responsiveness accounted for by changes in import penetration ratios from low wage countries.

Figure D3: General purpose technology share of high tech Manufacturing



Note: Tabulations from the LBD by authors. General purpose high tech 4-digit industries are NAICS 3341 (Computers), NAICS 3342 (Communication Equipment) and NAICS 3344 (Semi-conductors).

Figure D4: Change in responsiveness due to industry composition changes (high tech)



Note: Specification (2) as in Table 1 estimated for every 6-digit NAICS industry but without any trend effects. Reported coefficients are employment-weighted averages of the 6-digit NAICS industry estimated coefficients. Employment-weights vary by year.

Table D1: Employment growth, lagged productivity, and import penetration

TFP*Young	0.2085***
	(0.0390)
TFP*Young*Trend	0.0298***
	(0.0061)
TFP*Young*Trend^2	-0.0011***
	(0.0002)
TFP*Mature	0.1246***
	(0.0174)
TFP*Mature*Trend	0.0052**
	(0.0026)
TFP*Mature*Trend^2	-0.0003***
	(0.0001)
TFP*Young*Import Penetration	-0.0037***
	(0.0011)
TFP*Mature*Import Penetration	0.0002
	(0.0004)

Notes: Standard Errors in Parentheses. tech Sample is more than 120000 plant-year observations from 1981-2010. Young firms have age less than 5. Unreported are estimates of controls including year effects, state effects, firm age dummies, firm size dummies, log plant level employment in period t, state cyclical indicators (change in state level unemployment rate), state cyclical indicators interacted with TFP, and a main effect for the 6-digit import penetration ratio. All variables that use TFP including all interactions are fully interacted with firm age dummies. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.