

Changing Business Dynamism and Productivity: Shocks vs. Responsiveness

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

May 2019

Abstract

The pace of job reallocation has declined in the U.S. in recent decades. Existing literature suggests that job reallocation is productivity enhancing, so understanding the decline is important. We draw insight from canonical models of business dynamics in which reallocation can decline due to (a) lower dispersion of idiosyncratic shocks faced by businesses, or (b) weaker marginal responsiveness of businesses to shocks. We show that shock dispersion has actually risen, while the responsiveness of business-level employment to productivity has weakened. Using a novel model-inspired productivity counterfactual exercise, we find that weaker responsiveness accounts for a significant drag on aggregate productivity.

*Federal Reserve Board, University of Maryland and NBER, U.S. Census Bureau, and U.S. Census Bureau, respectively. John Haltiwanger is also a Schedule A part-time employee of the U.S. Census Bureau at the time of the writing of this paper. 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, three anonymous referees, 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, BYU, University of Chicago, Drexel University, the Federal Reserve Board, the Richmond Fed, 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. (2017). 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 motivating facts

Changing patterns of business dynamics—the entry, growth, decline, and exit of businesses—have attracted increasing attention in recent literature. In particular, since the early 1980s the U.S. has seen a decline in the pace of “business dynamism” across many measures including the rate of job and worker reallocation, the rate of internal migration, the rate of business entry, and the prevalence of high-growth firm outcomes.¹ Declining business dynamism has attracted attention in part because business dynamics are closely related to aggregate productivity growth in healthy market economies, reflecting the movement of resources from less-productive to more-productive uses (Hopenhayn and Rogerson (1993), Foster, Haltiwanger, and Krizan (2001)). In this paper, we investigate the declining pace of job reallocation specifically and examine its consequences for aggregate productivity. We draw insights from canonical models of firm dynamics that predict declining reallocation reflects a weakening of the responsiveness of individual businesses to their underlying productivity and profitability. We show that weaker responsiveness has significant negative implications for aggregate productivity.

Job reallocation measures the pace of job flows across businesses and is defined as total job creation by entering and expanding establishments plus total job destruction by downsizing and exiting establishments.² Figure 1 shows the pace of aggregate job reallocation for the U.S. overall, the manufacturing sector, and the high-tech sector. The U.S. experienced a decline in the pace of job reallocation since the early 1980s, even in the high-tech sector that saw a decline starting in the early 2000s. Understanding the causes of declining job reallocation has proven difficult. Decker et al. (2016b) show that it reflects in part a decline in firm-level growth rate skewness, or high-growth firm activity generally, but they do not investigate underlying causes. Young firms tend to exhibit a higher pace of job reallocation, and the share of activity accounted

¹ For job reallocation and employment volatility see Davis et al. (2007), Decker et al. (2014), and Decker et al. (2016b). For business entry see Decker et al. (2014) and Karahan, Pugsley, and Sahin (2018). For worker reallocation see Hyatt and Spletzer (2013) and Davis and Haltiwanger (2014). For migration see Molloy et al. (2016). For high-growth firms see Decker et al. (2016b), Haltiwanger, Hathaway, and Miranda (2014), and Guzman and Stern (2016).

² Job reallocation is typically expressed as a rate by dividing by total employment. The denominator in the job reallocation rate for year t is the average of total employment in years t and $t - 1$, following Davis, Haltiwanger, and Schuh (1996). An “establishment” is a specific business operating location, while a “firm” is collection of one or more establishments under common operational control.

for by young firms has declined (Decker et al. (2014)), so some decline in reallocation is to be expected given composition effects. However, most of the variation in reallocation rates in recent decades has occurred within narrow age classes.³

We study the causes of changing job reallocation rates motivated by the framework of standard models of firm dynamics following Hopenhayn (1992) and a rich subsequent literature. In such models, aggregate reallocation arises from businesses' responses to their constantly shifting individual productivity and profitability environment. Nascent and existing businesses facing strong idiosyncratic productivity and profitability conditions enter the market or expand (job creation), while those facing weak conditions downsize or exit (job destruction). The overall reallocation rate reflects the aggregation of these individual decisions. As such, a decline in the pace of reallocation can arise from one of two forces. First, the dispersion and volatility of idiosyncratic (business-level) conditions (which are called "shocks" in model terms) could decline; in other words, a more tranquil business environment could reduce the need and incentives for businesses to frequently and significantly change their size or operating status. Second, or alternatively, the business level *responsiveness* to those shocks could weaken; that is, businesses may hire or downsize less in response to a given shock, perhaps due to rising costs of factor adjustment.

These model-based considerations give rise to two competing hypotheses for declining job reallocation rates: the "shocks" hypothesis, in which the dispersion of idiosyncratic productivity or profitability realizations has declined; and the "responsiveness" hypothesis, in which businesses have become more sluggish in responding to realized shocks.⁴ Existing literature suggests that much of this responsiveness is productivity enhancing in the sense that it reflects movement of resources (labor in this case) from low-productivity to high-productivity businesses, increasing aggregate productivity through composition effects. Therefore, if responsiveness has declined, productivity growth could suffer as a result.⁵

³ Appendix III describes a shift-share exercise to study the role of composition effects across firm age for explaining the overall decline in job reallocation. Figure A3 in the appendix reports the results.

⁴ The general "shocks versus responsiveness" framework has proven useful elsewhere; see Berger and Vavra (forthcoming).

⁵ A one-time increase in adjustment frictions lowers the level of productivity in the steady state in a standard model. An increase in adjustment frictions over a period of time, along with resulting transition dynamics, can yield a decline in the productivity growth rate over an extended period of time.

Indeed, productivity growth has slowed in the U.S. in recent decades. Fernald (2014) and others document a surge in productivity growth in the late 1990s but a significant slowdown in the mid-2000s, led by technology-using and producing industries.⁶ At first glance the overall pattern of job reallocation, which shows a steady decline since the early 1980s, does not match up well with recent productivity dynamics. But the pace of reallocation in the critical high-tech sector, shown on Figure 1, mimics the pattern of aggregate productivity growth, rising in the 1990s before falling in the 2000s.

We first explore an off-the-shelf model of business dynamics to draw out critical insights about the “shocks” and “responsiveness” hypotheses. The simple model is purely illustrative but nevertheless yields rich empirical predictions that can be confronted with business microdata. We use the model to develop a novel framework for empirically assessing the implications of declining job reallocation for aggregate productivity. We then bring our hypotheses to business microdata for the U.S. We show that the dispersion of “shocks” faced by individual businesses has not in fact declined but has risen.⁷ However, business-level responsiveness to those shocks has declined markedly in the manufacturing sector and in the broader U.S. economy. Moreover, within high-tech industries, responsiveness rose during the 1980s and 1990s before declining thereafter, mimicking both the pattern of overall job reallocation in high-tech and the pattern of U.S. aggregate productivity dynamics.

Weaker responsiveness implies that productivity-based selection has weakened. We quantify the aggregate importance of changing business-level responsiveness using a novel counterfactual exercise motivated by our model. This allows us to isolate the effect of changing productivity responsiveness on aggregate productivity. We find that declining responsiveness is a significant “drag” on aggregate productivity; if the responsiveness of manufacturing establishments in the 2000s could immediately return to the strength of responsiveness seen in the early 1980s, we estimate that annual sector-wide TFP would initially increase by about 2 log points. We find similar orders of magnitude for the effects of weakening responsiveness on aggregate labor productivity for U.S. sectors broadly (beyond just manufacturing). While weaker responsiveness is likely not the dominant driver of slowing U.S. productivity growth, we

⁶ See also Byrne, Fernald, and Reinsdorf (2016); Gordon (2016); Syverson (2016).

⁷ Rising labor productivity dispersion outside manufacturing was first documented in a related working paper, Decker et al. (2016a) and in Barth et al. (2016). Andrews et al. (2015) documented rising gaps in the growth of firm-level labor productivity in several OECD countries.

find it is a nontrivial contributor (consistent with Decker et al. (2017), which instead relies on simple accounting decompositions).

Taken together, our results suggest that declining reallocation is not simply a benign result of a less turbulent economy. Rather, declining reallocation appears to reflect weaker responses of businesses to their own economic environment, and the consequences of weaker responsiveness for aggregate living standards are substantial due to the important role of productivity selection. Determining the causes of weakening responsiveness is beyond the scope of this paper; however, we describe several possible avenues of investigation that are suggested by standard models. In considering possible explanations, we emphasize that our finding of reduced responsiveness can alternatively be interpreted through the lens of a model of reduced-form wedges as in Hsieh and Klenow (2009). By reduced-form wedges, we mean anything that impedes firms from reaching their optimal frictionless size in response to changes in fundamentals like productivity. If such wedges are positively correlated with fundamentals, an increase in that correlation also yields a decline in responsiveness and an increase in revenue productivity dispersion.⁸ There are numerous possible sources of an increase in the correlation between reduced-form wedges and fundamentals; one source is an increase in adjustment frictions, but the broader perspective afforded by the Hsieh and Klenow (2009) model suggests a wider range of possible factors.

Section II describes our conceptual framework, which consists of a widely used model of business dynamics, and its empirical predictions for the “shocks” and “responsiveness” hypotheses; the section also describes a model-implied framework for assessing aggregate productivity implications. Section III describes our data, including our measures of productivity. Section IV describes our empirical approach and results on “shocks” and “responsiveness.” Section V quantifies the implications for aggregate productivity. Section VI describes robustness exercises, and section VII concludes.

II. Conceptual framework

To understand the secular trends in reallocation, we begin by using an off-the-shelf model with adjustment costs. We briefly describe the model and its empirical implications here; see

⁸ A positively correlated wedge means that the implicit tax on scale is positively correlated with fundamentals so that higher productivity businesses are smaller than optimal.

Appendix I for more detail. While the model is not novel, it is useful for illustrating the economics of job reallocation and motivating our empirical approach. We also emphasize that our empirical analysis can be interpreted through the lens of a more reduced-form wedge approach, as we explain below.

A. Basic model intuition

Consider a canonical model of firm dynamics with factor adjustment costs in the tradition of Hopenhayn and Rogerson (1993).⁹ For simplicity we abstract from firm entry and exit, and we include only one factor of production (labor). Firms face idiosyncratic productivity shocks, where the realization of productivity in the current period for firm j , A_{jt} , is drawn from a persistent AR(1) process. Labor adjustment—net hiring or downsizing—is subject to adjustment costs.¹⁰ Idiosyncratic productivity shocks should be interpreted broadly here as reflecting both technical efficiency and product appeal or demand-side idiosyncratic shocks. While we do not distinguish between these components, the conceptual and measurement framework we use below recognizes that firm-level prices are not directly measured and are likely endogenous.

The resulting decision rule for firms’ net hiring rates implicitly reflects adjustment costs and can be expressed as a net employment growth rate:

$$g_{jt} = f_t(A_{jt}, E_{jt-1}), \tag{1}$$

where the state variables are the productivity realization A_{jt} and initial employment E_{jt-1} , both of which are observed prior to the growth decision, and g_{jt} is employment growth from $t - 1$ to t . Observe that the policy function can be characterized as depending on the *level* of the productivity realization, not the change—a consequence of adjustment costs; see Cooper, Haltiwanger, and Willis (2007, 2015) and Elsby and Michaels (2013). This is convenient, since empirically it is easier to rely upon productivity levels than changes (as we explain below).

The basic intuition of the model follows. First, consider a version of the model with no adjustment frictions. Faced with revenue function curvature arising from decreasing returns to scale and/or imperfect competition, each firm has an optimal size (i.e., employment) determined entirely by A_{jt} . When a new period begins, each firm observes a new A_{jt} realization and, if $A_{jt} \neq A_{jt-1}$, hires or downsizes to reach the new optimal size implied by A_{jt} . The size of the

⁹ We use the term “firm” loosely in this section. Our empirical exercises use both firm-level and establishment-level data.

¹⁰ We use kinked adjustment costs, but our qualitative predictions are consistent with convex adjustment costs.

labor adjustment will depend on the magnitude of $(A_{jt} - A_{jt-1})$, which summarizes the difference in optimal size between the two productivity realizations.

The addition of adjustment frictions, however, makes optimal size depend on initial employment as well as the *level* of A_{jt} . Faced with costs on labor adjustment, firms no longer adjust their labor demand to reach the firm size that would be implied by A_{jt} in a frictionless world.¹¹ Instead, under standard assumptions, adjustment costs reduce the magnitude of adjustments, on average, and create dependence between employment growth and both the *level* of realized productivity and initial employment. The connection between growth and productivity levels arises because firms are never quite at the size they would prefer in a frictionless world and so are continually (or sporadically) adjusting toward that size. Given persistence in the A_{jt} process, high levels of A_{jt} will be associated with employment growth as firms gradually move toward the frictionless optimal, while low levels of A_{jt} will be associated with employment reductions. These dynamics are illustrated in Figure 1 of Cooper, Haltiwanger, and Willis (2007), which shows the relationship between A_{jt} and g_{jt} that emerges in a formal structural model of adjustment costs; the theoretical model in our Appendix I has the same predictions.

B. Shocks vs. responsiveness

This type of model suggests two possible explanations for declining job reallocation. First, since reallocation arises from the growth responses of individual firms to their individual productivity draws, a decline in the dispersion of productivity will reduce reallocation. We call this the “shocks” hypothesis since it arises from the nature of idiosyncratic shocks. The dispersion of shocks determines the distribution of sizes (employment) to which firms aspire; a decline in shock dispersion reduces the range of optimal sizes and therefore reduces the scope for employment growth in response to new shocks. Second, declining reallocation may reflect a dampening of the response of firms to their own productivity realizations. We call this the “responsiveness” hypothesis since it arises from the way in which firms choose to respond to their environment. In the canonical model, weaker responsiveness to shocks results from an increase in adjustment costs (or other changes we discuss below).

¹¹ This intuition holds for both convex and non-convex costs, as we discuss in the appendix.

We now describe these model implications in more detail. We measure “responsiveness” in the model in an empirically replicable way. Consider a regression of firm-level employment growth on firm-level productivity:

$$g_{jt+1} = \beta_0 + \beta_1 a_{jt} + \beta_2 e_{jt} + \varepsilon_{jt+1} \quad (2)$$

where we define $g_{jt+1} = (E_{t+1} - E_t) / (.5E_{t+1} + .5E_t)$, also known as the Davis, Haltiwanger, and Schuh (1995) growth rate (hereafter DHS), because the DHS growth rate is widely used in the literature and accommodates exit. The lowercase variables a and e refer to the logs of productivity and employment, respectively. The coefficient β_1 is a direct measure of the responsiveness of growth to productivity at the firm level.¹² More broadly, this growth regression can be thought of as an estimated linear approximation of the employment growth policy function given by equation (1). We run this regression on simulated data generated by model steady states for various parameterizations of the dispersion of shocks (the standard deviation of a_{jt}) and the magnitude of labor adjustment costs.

We conduct two experiments (comparing steady states) to illustrate the properties of the model. First, we explore the effects of changes in productivity dispersion. Figure 2a shows job reallocation, the productivity responsiveness coefficient β_1 , and the standard deviation of revenue per worker, each as a function of parameterized dispersion in a_{jt} . Note that in the absence of adjustment costs, equalization of marginal products implies zero dispersion of revenue per worker; with adjustment costs, revenue per worker is positively correlated with a_{jt} and exhibits positive dispersion.

Each of reallocation, responsiveness, and revenue productivity dispersion is increasing in a_{jt} dispersion; that is, as a_{jt} dispersion rises, the reallocation rate increases, measured responsiveness rises, and revenue productivity becomes more dispersed. Therefore, this model can generate a decline in job reallocation if a_{jt} dispersion falls; other symptoms of declining a_{jt} dispersion are weaker responsiveness and lower revenue productivity dispersion.¹³

¹² The regression timing convention shown in equation (2) is designed to match our empirical work and differs slightly from the timing in equation (1); the qualitative results are robust to other timing conventions.

¹³ We use kinked adjustment costs as our baseline case in the appendix. Kinked adjustment costs give rise to inaction ranges. As productivity dispersion falls there is a decrease in the fraction of firms that make zero adjustment (i.e., the “real options” effect). But declining productivity dispersion also implies smaller adjustments among those firms that do adjust (i.e., the “volatility” effect). Vavra (2014) argues that the volatility effect dominates the real options effect in the steady state, a general result extending back to Barro (1972). Bloom (2009),

Second, we explore the effects of changes in adjustment costs. Figure 2b shows the symptoms of an increase in the cost of employment downsizing from baseline levels: as the cost rises, reallocation falls, responsiveness weakens, and the dispersion of revenue productivity rises. Hence, the model can generate lower reallocation through higher adjustment costs, with additional symptoms of declining responsiveness and higher labor productivity dispersion.

The model therefore suggests simple empirical tests of the “shocks” vs. “responsiveness” hypotheses. The shocks hypothesis implies declining dispersion of both shocks and revenue productivity, along with declining marginal responsiveness. The responsiveness hypothesis implies declining reallocation, declining responsiveness, and *rising* revenue productivity dispersion. In our data, we can observe responsiveness (by estimating regressions analogous to equation (2)) as well as dispersion in shocks and revenue productivity. Given that marginal responsiveness changes with both changes in the distribution of shocks and adjustment frictions, examining responsiveness alone is not sufficient to distinguish between the changing shocks vs. adjustment frictions hypotheses. However, without a decline in the dispersion of idiosyncratic shocks, evidence of declining responsiveness requires increasing adjustment frictions (or some other rising wedge as discussed below). Moreover, examining changing responsiveness to TFP, changes in reallocation, and changes in revenue productivity dispersion using the same data and measurement approaches yields a joint test of the alternative hypotheses.

These predictions also depend on the persistence of shocks (as discussed in Cooper, Haltiwanger, and Willis (2007, 2015) and in our Appendix I). We find (empirically) that idiosyncratic productivity shocks are highly persistent, and we assume this in our model calibration. For a given level of adjustment costs, declining persistence can also yield a decline in reallocation and responsiveness: firms have little incentive to respond to low-persistence shocks. In our empirical analysis below, we present evidence that there has not been a significant change in the persistence of shocks. However, this discussion suggests it is also useful to distinguish between the realization of the shocks and the innovation to the shock. That is, assuming $a_{jt} = \rho a_{jt-1} + \eta_{jt}$, the responsiveness hypothesis concerns not only responsiveness to a_{jt} but also responsiveness to η_{jt} , which also declines with adjustment costs.

Bloom et al. (2018), and others use a similar model to study the effects of uncertainty on business cycles; even in their model, the volatility effect dominates *at annual frequency* (see also Bachmann and Bayer (2013)).

The gold standard empirical test of the responsiveness hypothesis is to estimate the changing relationship between the growth rate of employment and a_{jt} (and, in addition, the changing relationship between the growth rate of employment and η_{jt}). For the manufacturing sector, we can construct measures of a_{jt} and η_{jt} . For other sectors, we can only measure revenue per worker. However, the responsiveness hypothesis also implies a declining covariance between growth and the realization of revenue per worker. An increase in adjustment costs increases the correlation between a_{jt} and revenue per worker and, therefore, causes a decline in this covariance between employment growth and revenue per worker. Given this auxiliary prediction, we also explore changing “responsiveness” for non-manufacturing businesses using the changing relationship between employment growth and revenue per worker.

Our discussion thus far has focused on the intensive margin of responsiveness. However, related predictions apply for the extensive margin: Hopenhayn and Rogerson (1993) find that a rise in adjustment frictions reduces entry and exit. The empirical prediction of increased adjustment costs, then, is that not only will the growth of continuing firms become less responsive to firm productivity, but so will exit, a prediction we explore below.

Our motivating discussion also neglects post-entry dynamics from learning that can influence the responsiveness of both the extensive and intensive margins by firm age (see, e.g., Jovanovic (1982)). We consider this possibility in our empirical analysis. This variation by firm age is interesting in its own right but also permits us to abstract from changes in average responsiveness due to the changing age structure of firms. Given the decline in the U.S. firm entry rate in recent decades, if young firms have different average responsiveness from mature firms, aggregate responsiveness could have changed due to composition effects. We control for potentially exogenous changes in entry rates in the U.S. by studying empirical moments within firm age groups.

C. An increasingly correlated wedge interpretation

We have thus far motivated the shocks vs. responsiveness hypotheses in the context of canonical models with adjustment costs. However, our empirical focus is more general: we seek evidence on how responsiveness has changed over time for any reason. While adjustment costs are a natural way to generate changes in responsiveness, the changing shock vs. responsiveness hypotheses are robust to using a reduced-form wedge model. As an alternative to our main framework, in Appendix I we describe a “wedges” model in the spirit of Hsieh and Klenow

(2009). There we show that an increase in the correlation between wedges and fundamentals yields the same basic predictions as a rise in adjustment frictions: a decline in reallocation, a decrease in responsiveness to fundamentals, and a rise in revenue productivity dispersion. With correlated wedges, a decline in the dispersion of productivity shocks also yields a decline in reallocation and revenue productivity dispersion.

This wedge interpretation could be viewed as a reduced form that encompasses the adjustment cost friction interpretation discussed above (albeit with some important subtle differences given the explicitly dynamic components of an adjustment cost model). But this interpretation also may capture other possible increases in correlated wedges. For example, rising dispersion in variable markups that are correlated with fundamentals can play a similar role (see, e.g., De Loecker, Eeckhout, and Unger (2018)). Our purpose is to study changing patterns of job reallocation by documenting changes in business-level responsiveness; identifying the source of the increase in correlated wedges is beyond the scope of this paper.¹⁴

D. Implications for aggregate productivity

Finally, we use the model to provide an empirical framework for quantifying the importance of changing responsiveness. We show in Figure 2c and Appendix I that weaker productivity responsiveness arising from increased adjustment costs (or other reasons) implies a weakening of productivity selection, with negative implications for aggregate productivity (which is aggregate output per worker in our model).¹⁵ Given the many factors that can affect aggregate productivity over time, we seek a method for isolating the effect of changing responsiveness on aggregate productivity growth.

We use a novel diff-in-diff counterfactual exploiting our changing responsiveness regressions to isolate the impact of changing responsiveness. Consider two different regimes: high responsiveness (in which β_1 from equation (2) is high, or $\beta_1 = \beta_1^H$) and low responsiveness ($\beta_1 = \beta_1^L$, with $\beta_1^H > \beta_1^L$). In the model analysis in Appendix I, this corresponds to low and high adjustment costs, respectively. Holding constant the distribution of productivity a_{jt} and employment in period t , these two regimes result in different predicted employment shares

¹⁴ Investigating the potential role of changing markups for changing responsiveness requires additional research. Amongst other things, rising markups implies that our revenue elasticity estimate should reflect those changes. One exercise that suggests that our results are robust to this concern is that we have (in unreported exercises) estimated the TFPS using the revenue cost share approach using Divisia-based (average of $t - 1$ and t shares) and obtained similar results.

¹⁵ Hopenhayn and Rogerson (1993) show the same finding from an increase in adjustment frictions.

across firms in period $t + 1$. In the high responsiveness regime, the high value of β_1^H means that high-productivity firms grow more rapidly, gaining employment share, while this effect is weaker in the low responsiveness regime. Given the time- t distribution of employment and productivity, let the time- $t + 1$ employment shares resulting from β_1^H be given by $\hat{\theta}_{jt+1}^H$, and let the employment shares resulting from β_1^L be given by $\hat{\theta}_{jt+1}^L$. The high-responsiveness and low-responsiveness regimes generate two different implied changes in the weighted average of firm-level productivity implied by changes in shares alone. These are given by:

$$\sum_j \hat{\theta}_{jt+1}^H a_{jt} - \sum_j \theta_{jt} a_{jt}, \quad (3)$$

$$\sum_j \hat{\theta}_{jt+1}^L a_{jt} - \sum_j \theta_{jt} a_{jt}, \quad (4)$$

where θ_{jt} corresponds to the initial (true) distribution of employment shares, and a_{jt} corresponds to the initial distribution of productivity. Both distributions are held constant across these counterfactuals. Expressions (3) and (4) are thus the year-over-year changes in the weighted average of firm-level productivity implied by model-predicted changes in employment shares in the high and low responsiveness regimes, respectively. We take the difference in these two counterfactuals to construct:

$$\Delta_t^{t+1} = \sum_j \hat{\theta}_{jt+1}^L a_{jt} - \sum_j \hat{\theta}_{jt+1}^H a_{jt}, \quad (5)$$

Therefore, Δ_t^{t+1} is the implied difference in the change in the weighted average of firm-level productivity between t and $t + 1$ due to differential responsiveness, holding constant the actual distribution of firm-level productivity.

The diff-in-diff counterfactual Δ_t^{t+1} closely approximates the change in aggregate productivity caused by a change in adjustment costs. Figure 2c reports true aggregate productivity relative to the baseline model for various adjustment cost scenarios (solid line). The dashed blue line refers to the diff-in-diff construct (equation (5)) for various “high” adjustment cost scenarios relative to the “low” baseline, where the regression model predicting $\hat{\theta}_{jt+1}$ (equation (2)) relies on the concurrent value of firm productivity a_{jt+1} ; and the dashed orange line is the same object but relies on *lagged* productivity a_{jt} (that is, lagged relative to $\hat{\theta}_{jt+1}$) for predicting $\hat{\theta}_{jt+1}$ in the regression model. While the concurrent productivity specification is theoretically preferred, as we discuss below, our *empirical* model timing lies somewhere between the concurrent and the lag specifications. In either case, the diff-in-diff productivity

implications of rising adjustment costs are similar to the true productivity effects, with the true effect lying between two diff-in-diff objects.

The intuition for this close approximation of the true productivity effect is that this counterfactual holds the distribution of firm-level state variables (i.e., productivity and employment in t) constant and isolates the impact of differences in the pace of reallocation across responsiveness regimes. Lower responsiveness yields lower counterfactual changes in productivity. This mimics the aggregate productivity change given that the responsiveness regressions approximate the growth policy rules at the firm level and the counterfactual holds everything constant except adjustment cost-induced changes in responsiveness.¹⁶ In our empirical analysis, Δ_t^{t+1} is readily computable from regressions estimating changing responsiveness along with the underlying true (observed) distributions of firm-level productivity and employment. Based on the above, we interpret Δ_t^{t+1} as quantifying (and isolating) the effect of changing responsiveness on the difference in aggregate productivity between t and $t + 1$.

III. Data and measurement

The main database for our analysis is the U.S. Census Bureau's Longitudinal Business Database (LBD), to which we attach other data as detailed below. The LBD includes annual location, employment, industry, and longitudinal linkages for the universe of private non-farm establishments, with firm identifiers based on operational control (not an arbitrary tax identifier).¹⁷ Employment measures in the LBD come from payroll tax and survey data. We use the LBD for 1981-2013 (during which consistent establishment NAICS codes are available from Fort and Klimek (2016)). For some exercises we focus on the high-tech sector; we define high-tech on a NAICS basis following Hecker (2005).¹⁸ As in previous literature, we construct firm age as the age of the firm's oldest establishment when the firm identifier first appears in the data, after which the firm ages naturally.

Importantly, for both our manufacturing and private sector economy analysis, we use the LBD to measure employment growth, initial employment, and exit (characterized as an

¹⁶ In the prior working paper version (Decker et al. (2018)), we showed that this counterfactual tracks changes in aggregate productivity much closer than using an Olley and Pakes (1996) decomposition.

¹⁷ See Jarmin and Miranda (2002) for a full description of the LBD.

¹⁸ Hecker (2005) defines industries as high-tech based on the 14 four-digit NAICS industries with the largest share of STEM workers. This definition includes industries in manufacturing (NAICS 3254, 3341, 3342, 3344, 3345, 3364), information (5112, 5161, 5179, 5181, 5182), and services (5413, 5415, 5417).

establishment or firm that has positive employment activity in t and zero activity in $t + 1$). We use these LBD measures of growth and exit even when we merge in productivity measures from manufacturing (described next). This enables us to overcome some of the limitations of the ASM/CM data for panel analysis, as we clarify below.

A. Manufacturing: Measuring establishment-level productivity

We construct establishment-level productivity for over 2 million plant-year observations (1981-2013) using updated data following the measurement methodology of Foster, Grim, and Haltiwanger (2016) (hereafter FGH) combining the Annual Survey of Manufacturers (ASM) with the quinquennial Census of Manufacturers (CM); see Appendix II for detail. The resulting ASM-CM is representative of the manufacturing sector in any given year, but it is based on a rotating sample and thus lacks the complete longitudinal coverage of the LBD. To compensate, we integrate the ASM-CM into the LBD to obtain establishment-level employment growth.¹⁹ Thus, a critical feature of our empirical approach (for manufacturing) is integrating the high-quality longitudinal growth measures from the LBD in any given year with the cross-sectional measures of productivity (broadly defined) from the ASM-CM.

The productivity shocks we measure are intended to capture changes in both technical efficiency and demand or product appeal. To make our measurement approach transparent, it is helpful to be explicit about the assumed production and demand structure.²⁰ Consider establishment-level demand function $P_{jt} = D_{jt}Q_{jt}^{\phi-1}$ (where D_{jt} is an idiosyncratic demand shock, $\phi - 1$ is the inverse demand elasticity, and j indexes establishments) with Cobb-Douglas production, that is, $Q_{jt} = \tilde{A}_{jt} \prod X_{jt}^{\alpha_x}$ for inputs X_{jt} (where \tilde{A}_{jt} is technical efficiency, or TFPQ). A composite measure of productivity “TFP” reflecting idiosyncratic technical efficiency and demand shocks can be defined as $A_{jt} = D_{jt}\tilde{A}_{jt}^{\phi}$. The ASM/CM data provide survey-based

¹⁹ We also use propensity score weights (based on a logit model of 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). 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.

²⁰ Here we use notation that is the same as in our model description in Appendix I.

measures of revenue, capital (K), labor hours (L), materials (M), and energy (E).²¹ Then establishment revenue is given by (lower case variables are in logs):

$$p_{jt} + q_{jt} = \beta_k k_{jt} + \beta_l l_{jt} + \beta_m m_{jt} + \beta_e e_{jt} + \phi \tilde{a}_{jt} + d_{et}, \quad (6)$$

where $\beta_x = \phi \alpha_x$ for factor X , and t denotes time (in years).²² The β_x coefficients are factor revenue elasticities that reflect both demand parameters and production function factor elasticities. The implied *revenue function residual*, which we denote as TFP, is given by:

$$TFP_{jt} = p_{jt} + q_{jt} - (\beta_k k_{jt} + \beta_l l_{jt} + \beta_m m_{jt} + \beta_e e_{jt}) = \phi \tilde{a}_{jt} + d_{jt}, \quad (7)$$

that is, this measure of TFP is a composite of idiosyncratic technical efficiency and demand *shocks*. In terms of the conceptual framework described previously (and in Appendix I), this is the relevant measure of fundamental shocks. With estimates of the revenue elasticities, this measure of TFP can be computed from observable establishment-level revenue and input data. We refer to this measure as “TFP” or “productivity” in what follows, but it should be viewed as the composite shock reflecting both technical efficiency and product demand or appeal. Our use of the revenue function residual to capture fundamentals is not novel to this paper. Cooper and Haltiwanger (2006) estimate the revenue function residual in their analysis of capital adjustment costs. Hsieh and Klenow (2009) use a closely related measure as their empirical measure of “TFPQ.”²³ Blackwood et al. (2019) use a similar measure in their analysis of allocative efficiency as a proxy for “TFPQ.”²⁴

²¹ Labor input is total hours measured from the survey responses in the ASM/CM. We estimate factor elasticities for equipment and structures separately but refer only to generic “capital” for expositional simplicity here. See Appendix II for more discussion of production factor measurement in the data.

²² Output (q) is total value of shipments plus total change in the value of inventories, deflated by 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 (each deflated by an industry-level deflator). Outputs and inputs are measured in constant 1997 dollars. More details are in Appendix II.

²³ The empirical measure of TFPQ used by Hsieh and Klenow is proportional to the revenue function residual measure of TFP. That is, the measure they use for TFPQ (in logs) is equal to our measure of TFP divided by ϕ . While proportional, it is more challenging to construct their measure of TFPQ since it also requires an estimate of ϕ . That is, it requires decomposing the revenue elasticities into their demand and output elasticities components; see Blackwood et al. (2019). Foster, Haltiwanger, and Syverson (2008) define TFPQ to be technical efficiency. The Hsieh and Klenow empirical measure is inclusive of any idiosyncratic demand shocks, as is our measure.

²⁴ The gold standard is to use establishment- or firm-level prices permitting separation of technical efficiency and demand (and also alternative estimate approaches for output and demand elasticities). However, such prices are available for the only limited products in the Economic Censuses (see Foster, Haltiwanger, and Syverson (2008)).

Below we discuss two alternative approaches to estimating the revenue function residual concept for TFP in (7), but first we describe another productivity concept that is widely used in the literature, “TFPR,” which is given by:

$$TFPR_{jt} = p_{jt} + q_{jt} - (\alpha_k k_{jt} + \alpha_l l_{jt} + \alpha_m m_{jt} + \alpha_e e_{jt}) = p_{jt} + \tilde{a}_{jt}, \quad (8)$$

The key conceptual and measurement distinction between TFP in (7) and TFPR in (8) is using revenue versus output elasticities. TFPR confounds technical efficiency and endogenous price factors. As emphasized by Foster, Haltiwanger, and Syverson (2008), Hsieh and Klenow (2009), and Blackwood et al. (2019), this implies it is an endogenous measure when prices are idiosyncratic and endogenous. Without frictions or wedges, TFPR will exhibit no within-industry dispersion and is therefore not an appropriate measure of fundamentals. With adjustment costs or correlated distortions, however, TFPR will be positively correlated with fundamentals. Empirically, the evidence shows that TFPR and fundamentals are strongly positively correlated (Foster, Haltiwanger, and Syverson (2008) and Blackwood et al. (2019)). The high correlation in practice helps rationalize the widespread use of TFPR as a measure of TFP in the empirical literature.²⁵ For our purposes, TFPR is a useful measure since in our model an increase in adjustment costs or correlated wedges yields a decline in the responsiveness of growth to TFPR and a rise in dispersion of TFPR. In this respect, TFPR has properties similar to revenue per worker. We emphasize that, given the endogeneity limitation of TFPR, we do not consider it to be a clean measure of “shocks” as are the TFP measures from (7). Rather, TFPR is a measure of revenue productivity (reflecting the product of prices and technical efficiency).

We now describe how we estimate our various manufacturing productivity measures. The construction of TFP from (7) requires estimates from of the β_x revenue elasticities. We obtain estimates in two different ways, resulting in two alternative TFP-based measures. Our first and preferred TFP measure relies on the first-order condition (for factor X) from static profit maximization:

$$\alpha_x \phi = \beta_x = \frac{W_{xt} X_{jt}}{P_{jt} Q_{jt}}, \quad (9)$$

where W_{xt} is the price of factor X such that β_x is the share of the factor’s costs in total revenue. The condition in (9) will not hold for all establishments at all times if there are adjustment frictions or wedges, but we only need (9) to hold on average when pooled through time and over

²⁵ It is also a measure of fundamentals if plants are price takers.

establishments within industries, an assumption commonly used in the literature (Syverson (2011)). We obtain factor shares of revenue from the NBER-CES database (at the 4-digit SIC level prior to 1997 and the 6-digit NAICS level thereafter) then extract revenue function residuals using equations (6) and (7). We call this measure TFPS (for “TFP-Shares”).

Our second TFP measure is based on estimation of the revenue function in (7) using the proxy method GMM approach of Wooldridge (2009), allowing elasticities to vary at the 3-digit NAICS level (see Appendix II for details; see other applications in, e.g., Gopinath et al. (2017) and Blackwood et al. (2019)). We refer to this measure as TFPP (for “TFP-Proxy”). The TFPP method allows us to avoid reliance on first-order conditions, but the estimation process involves high-order polynomials and so requires large samples.²⁶ Following the literature, then, we use higher levels of aggregation for estimating industry elasticities. We use 3-digit NAICS compared to the 6-digit NAICS used for TFPS. This limitation of the proxy methods makes TFPS our preferred measure, but our results are robust to using TFPP.

For the TFPR measure from (8), we construct output elasticities as cost shares of inputs out of total costs (see, e.g., Baily, Hulten, and Campbell (1992), Foster, Haltiwanger, and Krizan (2001), Syverson (2011), Ilut, Kehrig, and Schneider (2018), Bloom et al. (2018)).²⁷ Following related literature, we use the NBER-CES productivity database to recover factor cost shares.²⁸ Cost shares equal factor elasticities under the assumptions of cost minimization and full adjustment of factors; again, however, one need not assume full adjustment for each establishment in each time period but rather that this holds approximately when pooling across all plants in the same industry over time. Like TFPS, our TFPR measure avoids the noisiness of estimation and allows us to use output elasticities that vary only at the detailed industry level.

In all of our empirical work (whether using TFPS, TFPP, or TFPR), we deviate the (log of the) measure from its detailed industry-by-year mean. Our data are not ideally suited for tracking the persistence of and innovations to the TFP measures given the panel rotation of the ASM and our use of CM data. However, we overcome these issues by conducting analysis of

²⁶ Blackwood et al. (2019) find that the Wooldridge (2009) method residuals are highly sensitive to outliers but that pooling across more observations mitigates this problem. This is the reason we vary elasticities at the 3-digit NAICS level in this measure.

²⁷ Using cost shares of total costs requires an estimate or assumption of returns to scale. The standard approach for constructing TFPR is to assume constant returns to scale.

²⁸ We construct time-invariant elasticities; in unreported results, we allow elasticities to vary over time with a Divisia index and find similar results.

persistence and innovations by excluding years for which we do not have a representative sample of continuing plants in t and $t - 1$ in our ASM-CM data (first panel years and Census years). Finally, we acknowledge that our establishment-level measures of TFP are vulnerable to errors arising from omitted factors. In particular, use of intangible capital in production is picked up as productivity in our measures, a limitation to which other TFP measurement attempts in the literature are also subject. In the cross section, if establishments with more intangible capital are more likely to grow, we will observe that as high responsiveness to productivity.

B. Total economy revenue labor productivity

While TFP is the preferred concept in a shocks vs. responsiveness framework, we can only estimate TFP in the manufacturing sector. For the economy generally, we rely on revenue per worker (“revenue labor productivity” or RLP), which is necessarily a firm-level (rather than establishment-level) concept in our data. As discussed above, rising adjustment frictions or correlated distortions also implies rising RLP dispersion and declining “responsiveness” of growth with respect to RLP.

Combining LBD employment (collapsed from the establishment to the firm level) with revenue measures in the Census Bureau’s Business Register (BR) (aggregated across EIN reporting units to the firm level) yields an enhanced LBD that we refer to as the RE-LBD. Revenue data are available from 1996 to 2013 and are derived from business tax returns.²⁹ 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.³⁰ Merger activity is interesting but is outside the scope of our investigation, as it relies on productivity in multiple businesses and does not easily map into our model framework.

²⁹ See Appendix II for more details on revenue data construction. 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 Census considers part of that firm for employment purposes. We address potential match-driven selection bias by constructing inverse propensity score weights.

³⁰ The organic growth rate calculation is straightforward but requires highly specific definitions of firm-level employment. For a firm j , let E_{jt} be the sum of employment in year t among all establishments owned by firm j in year t , and let E_{jt-1} be the sum of employment in year $t - 1$ among all establishments owned by firm j in year t inclusive of establishments that closed between $t - 1$ and t . Then the firm-level growth rate is given by $g_{jt} = (E_{jt} - E_{jt-1}) / (0.5E_{jt} + 0.5E_{jt-1})$. See Haltiwanger, Jarmin, and Miranda (2013) for more discussion of organic firm growth.

Similar to our TFP construction, for RLP we use (log) revenue per worker deviated from detailed (6-digit NAICS) industry-by-year mean. We thereby control for 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 TFPR.

For firm-level exercises, we must assign firms an industry code. Following much of the literature, we assign each firm a consistent (non-time-varying) “modal” industry code based on the NAICS industry in which it has the most employment over time. In exercises reported in the earlier working paper version (Decker et al. (2018)), we show our results are robust to an alternative approach in which we explicitly control for all industries in which firms have activity rather than assigning each firm a single industry code. 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.

IV. Empirical approach and results

A. “Shocks” hypothesis

We now study the dispersion of our various productivity measures. Again, in each case we use *within-industry* productivity: for any productivity measure z (which is in logs), we specify establishment- or firm-level productivity as $z_{jt} - \bar{z}_t$, where \bar{z}_t is the average for plant j 's industry. Figure 3a reports the standard deviation of our three (within-industry, log) measures for manufacturing —TFPS, TFPP, and TFPR—averaged for the 1980s, the 1990s, and the 2000s (up through 2013). Our preferred measure, TFPS, sees an increase from about 0.46 in the 1980s to 0.51 in the 2000s. The other measures also show widening dispersion. Figure 3b reports the standard deviation of (within-industry, log) revenue labor productivity (RLP) for the total U.S. economy (the first column). Since our RLP data cover a shorter time span than our TFP data, we show more time detail. As is apparent, RLP dispersion has risen over this time period for the whole economy, showing that rising productivity dispersion is not just a manufacturing phenomenon. The remaining bars on Figure 3b report RLP dispersion for manufacturing only (for ease of comparison with Figure 3a); the second set of bars is RLP dispersion in the ASM

(which, along with the CM, is the source of our revenue TFP data), and the third set of bars is RLP dispersion in manufacturing from the RE-LBD.

Figures 3a and 3b reveal several insights. First, consistent with previous literature (e.g., Syverson (2004, 2011)), within-industry dispersion in TFP is large; for example, a level of 0.51 (51 log points) for TFPS implies that an establishment one standard deviation above the mean for its industry is about $e^{.51} \approx 1.7$ times as productive as the mean. Within-industry RLP is even more dispersed—as may be expected given potential dispersion in non-labor production factors, especially capital. Second, the three TFP measures, while substantially different in construction, yield broadly similar dispersion trends. Third, the rise in revenue productivity dispersion observed in manufacturing survey data is confirmed by administrative data (compare the second and third sets of bars on Figure 3b). Bils, Klenow, and Ruane (2017) argue that rising revenue productivity dispersion observed in the ASM is due to increasing survey-based measurement error, but Figure 3b shows that the rise in various measures of productivity dispersion in the U.S. is evident in administrative data, apparently not an artifact of survey limitations.³¹

In Figure 3c we report the dispersion of revenue TFP *innovations* (bottom left panel of Figure 3), and Figure 3d reports the *persistence* of revenue TFP levels (bottom right panel).³² The dispersion of innovations has also risen, while the persistence of shocks has declined only modestly, in recent decades.

Figure 3 implies that shock dispersion has not declined, as might be expected from declining reallocation, but, if anything, has actually risen. In other words, the dispersion and volatility of shocks faced by businesses have not evolved in a way that could explain declining job reallocation. The business environment has not become less idiosyncratically turbulent; if anything, it has become more so. The findings for the dispersion in the realizations of TFPS and TFPP and innovations are direct evidence of rising shock volatility. The findings for TFPR and RLP are indirect evidence. All else equal, the data on shock dispersion should imply a *rising* pace of reallocation, while we observe the opposite. We therefore reject the “shocks” hypothesis and now evaluate the “responsiveness” hypothesis.

³¹ Recall that while the ASM measures of revenue and employment are from survey responses, the RE-LBD measures are from business tax returns (for revenue) and payroll tax records (for employment).

³² The AR(1) estimates in Figure 3c for the TFPS and TFPP are somewhat lower than those in the literature (e.g., Foster, Haltiwanger, and Syverson (2008), which use a narrow sample of products, and Cooper and Haltiwanger (2006), which use only plants that produce from 1972-1988).

B. “Responsiveness” hypothesis: Initial exploration

We next evaluate the “responsiveness” hypothesis for declining job reallocation—that is, the hypothesis that declining job reallocation is a result of dampened responsiveness of firms and establishments to their idiosyncratic productivity shocks. The evidence of rising revenue productivity dispersion we document above already is consistent with responsiveness weakening; as shown in our model discussion, rising revenue productivity dispersion may reflect rising adjustment costs, changes in wedges, or rising dispersion of fundamentals. However, we can more directly test the responsiveness hypothesis by estimating responsiveness itself in the data.

We proceed in a manner analogous to our measurement of responsiveness in model-simulated data above; that is, we estimate an expanded version of equation (2):

$$g_{jt+1} = \beta_0 + \beta_1 a_{jt} + T(a_{jt}, t) + \beta_2 e_{jt} + T(e_{jt}, t) + X'_{jt} \Theta + \varepsilon_{jt+1}. \quad (10)$$

Equation (10) forms the core of our approach to measuring changes in responsiveness over time, so we will describe it in some detail. Individual establishments or firms are indexed by j , and time (in years) is indexed by t . The dependent variable, g_{jt+1} , is annual DHS employment growth between years t and $t + 1$.

In our baseline results, we measure productivity (a_{jt}) by the log of TFP.³³ We extend that baseline specification in a variety of ways, including the use of innovations to (rather than levels of) TFP. For our baseline specifications using the log of TFPS/TFPP (either realizations or innovations), β_1 estimates “responsiveness” (or the response of growth to productivity at the establishment or firm level) and corresponds to β_1 from equation (2), our responsiveness regression on model-simulated data. For extended analyses using TFPR and RLP, we obtain insights into changing responsiveness with respect to these measures. For all of our measures of productivity, we permit this responsiveness to vary over time via $T(a_{jt}, t)$ as described below.

Initial employment, another critical state variable in our model, is given by e_{jt} , which is measured as log establishment-level employment from the LBD. X'_{jt} includes detailed industry (e.g., 6 digit NAICS) interacted with year effects, establishment size (in the case of specifications for manufacturing), firm size, state fixed effects, the change in state unemployment rates (to

³³ Employment growth is measured from the LBD using employment for the payroll period covering March 12 in period t to $t + 1$. Productivity (e.g., TFPS) is measured for the calendar year t . Thus, the empirical timing of the data is close to the timing of the model-based specification in equation (2), though in Appendix I we show that the specific timing is not important for our main empirical implications.

measure state-level business cycle effects), and interaction terms between the change in state unemployment rates and productivity; our liberal inclusion of cyclical indicators is intended in part to avoid result contamination from the Great Recession.³⁴ We also include a complete set of detailed industry-by-year effects. We estimate equation (10) on our manufacturing establishment sample (covering 1981-2013) for our TFPS, TFPP, and TFPR measures, and on our total economy firm sample (covering 1997-2013) in which a_{jt} is replaced with the log of revenue labor productivity (RLP).

Equation (10) also allows productivity responsiveness to vary over time via $T(a_{jt}, t)$, which we define variously as follows:

$$T(a_{jt}, t) \in \begin{cases} \delta a_{jt} Trend_t, \\ \gamma_{97} a_{jt} \mathbb{I}_{\{t \geq 1997\}}, \\ \lambda_{80s} a_{jt} \mathbb{I}_{\{t \in (1980, 1990)\}} + \lambda_{90s} a_{jt} \mathbb{I}_{\{t \in [1990, 2000)\}} \\ + \lambda_{00s} a_{jt} \mathbb{I}_{\{t \geq 2000\}} - \beta_1 a_{jt} \end{cases} \quad (11)$$

The first element of (11) defines the time function as a simple linear trend with coefficient δ . The second element uses a dummy variable to split the manufacturing sample roughly in half, such that overall responsiveness is equal to β_1 prior to 1997 and $\beta_1 + \gamma_{97}$ thereafter. The third element allows responsiveness to vary by decade, where the final “decade” is 2000-2013; by subtracting $\beta_1 a_{jt}$, we remove the main effect specified in (10) so the decade coefficients can be interpreted in a fully saturated manner. We also permit the effects of initial employment to vary over time in an analogous fashion.

The specification given by (10) is analogous to our model-based exercise from equation (2) and produces a reduced-form yet direct estimate of policy functions generated by standard models. Moreover, by using DHS growth rates we can incorporate both the intensive margin and the extensive margin (exit) of plant- and firm-level growth.

It is worth emphasizing that the employment growth measure and the initial period t employment measure are from the LBD (not the ASM-CM); this is important for two reasons. First, the LBD growth measure uses longitudinal linkages available for all establishments. This

³⁴ In unreported exercises, we omit the cyclical controls in X'_{jt} and find very similar results. Moreover, to further ensure the Great Recession does not drive our results, in unreported exercises we end our sample in 2007 and still find very similar results.

means we can track employment growth from t to $t + 1$ for each establishment in the representative ASM-CM cross section for which we have TFP measures in t . When we use innovations to TFP we reduce the set of years available but, again, track the employment growth for all establishments for which we measure innovations in t . Second, the administrative measures of employment in the LBD used to measure growth and initial employment are high quality, reducing concerns about possible division bias from measurement error in initial employment. For the manufacturing analysis, the employment measure used to construct the growth rates and initial employment differs from the source data for total hours used to construct the TFP measures.³⁵ See Section VI.B below for further discussion and robustness analysis.

Table 1 reports results from the regression in (10).³⁶ Panel A, the top panel, reports establishment- and firm-level regressions using annual DHS employment growth (inclusive of exit) as the dependent variable, as in equation (10). Panel B reports the same regressions except using exit as the dependent variable in a linear probability model.³⁷ All regressions include the $X'_{jt}\Theta$ term from equation 10, but we do not report those coefficients.³⁸ We divide each panel into four parts reflecting our four productivity concepts: TFPS (in which factor elasticities are revenue shares), TFPP (in which factor elasticities are estimated with proxy method), TFPR (in which factor elasticities are simply cost shares), and RLP (real revenue per worker).

Consider the first section of Panel A, under the heading “TFPS (revenue share based)”. This section refers to establishment-level regressions in which the dependent variable is employment growth and the productivity variable a_{jt} is TFPS. The first column specifies changing responsiveness with the linear time trend described in (11). For TFPS, we estimate a base responsiveness coefficient of $\hat{\beta}_1 = 0.2965$, a significant positive number indicating strong selection early in the data time period, but we also find $\hat{\delta} = -0.0035$, which indicates responsiveness has weakened over time, as hypothesized. The regression reported in the next column uses the post-1997 responsiveness shifter from (11). Here we find a pre-1997 responsiveness coefficient of $\hat{\beta}_1 = 0.2905$, but after 1997 the responsiveness coefficient is equal to the base estimate plus the coefficient on the post-1997 interaction, $\hat{\gamma}_{97} - 0.0952$, for a total

³⁵ These features reduce concerns of division bias from measurement error in employment.

³⁶ All coefficients on Table 1 are statistically significant with $p < 0.01$.

³⁷ The exit specifications eliminate any concerns about division bias.

³⁸ We report only β_1 and the time function $T(a_{jt}, t)$ coefficients to satisfy data disclosure constraints.

responsiveness coefficient in the post-1997 period of 0.1953—a number that is still consistent with positive responsiveness and productivity selection, but much weaker responsiveness than in the earlier period. The next column reports estimates from the fully saturated decade indicators (λ) from (11). Here we can see the clear step down in productivity responsiveness, from about 0.29 in the 1980s to 0.20 in the 2000s, and the lower rows of this regression column report p values from t tests of equality between the various decade coefficients; in the case of TFPS, each decade coefficient is statistically different from the others.

The other sections of Panel A proceed analogously to the TFPS analysis, substituting the alternative productivity measures into otherwise identical regressions. Within manufacturing, while the quantitative results differ some between the alternative measures, and the exact timing of the decline in responsiveness varies somewhat, overall the qualitative results are strikingly similar and confirm a multi-decade decline in responsiveness.

The last section in Panel A refers to results for the whole economy using RLP (revenue labor productivity) as the productivity concept. Again, these data have a shorter time window. We estimate a clear decline in responsiveness, as shown by the negative, significant value for $\hat{\delta}$ (the trend term). The weakening of responsiveness we observe in manufacturing has occurred across the economy generally.

Panel B reports regressions that mimic those in Panel A except that the dependent variable is now exit (firm failure, a binary indicator that is unity if the firm exits the data between years t and $t + 1$) rather than growth. While the DHS growth rate indicator used in Panel A is inclusive of exit, it is useful to focus on the extensive margin in isolation. As noted in the theory discussion, canonical models imply that a rise in adjustment frictions not only reduces the responsiveness of growth but also reduces exit. In this case, a negative coefficient implies positive selection: businesses with high productivity draws are more likely to survive. The results of these regressions are similar to the growth regressions. We focus on our preferred measure, TFPS. The second column shows an exit coefficient that goes from -0.0801 in the pre-1997 period to -0.0340 thereafter. The third column shows a substantial and statistically significant weakening of exit responsiveness from the 1980s to the 1990s along with some further modest (and marginally significant) weakening in the 2000s. Intuitively, in the later years of the sample there are surviving establishments or firms whose relative productivity is so low that it would have prompted exit in earlier years.

The coefficients on Panel A reflect a linear relationship between growth—in percentage points—and the log of productivity, and the coefficients on Panel B reflect a linear relationship between productivity and exit probabilities. A useful way to quantify the magnitude of these coefficients—and of the decline in responsiveness—is to link them to the actual distribution of productivity. Since productivity is specified in log terms and deviated from its industry-year mean, we can easily express the coefficients as the difference in employment growth between the establishment (or firm) that is one standard deviation above its industry mean and the establishment (or firm) that is at the industry mean by multiplying each regression coefficient by the standard deviation of productivity.³⁹ Figure 4a reports the standard deviation of TFPS by decade; we take the average across decades (0.48) to isolate the effect of changing responsiveness (i.e., avoid confounding the responsiveness change with changes in dispersion) and multiply it by the decade coefficients found in the TFPS regressions on Table 1.

The result is on Figure 4a, where we flip the sign of the exit coefficients for comparability. During the 1980s, an establishment that was one standard deviation above its industry in terms of TFPS grew its employment (over one year) by 14 percentage points more than the industry mean, a striking illustration of the intensity of productivity selection within industries. That same establishment also faced an exit risk 3.7 percentage points lower than its industry mean. In the 1990s, the growth rate differential fell to 12 percentage points while the exit risk differential narrowed to 2.8 percentage points. By the 2000s, the growth differential was 10 percentage points and the exit risk differential was 2.5 percentage points. While productivity selection is still clearly evident, the decline in responsiveness has weakened selection materially, substantially narrowing the growth and survival advantage of high-productivity establishments.

Figure 4b shows the same differentials for RLP (using the standard deviation of RLP from Figure 3b); since we do not have decade dummy coefficients for RLP, here we use $\hat{\delta}$, the linear trend coefficient, combined with the base coefficient β_1 , to construct annual responsiveness coefficients, then we report multi-year averages at the beginning and end of the period. The result is similar: the growth and survival advantage of high-productivity firms (those firms whose revenue per worker is one standard deviation above their industry mean), while still

³⁹ Note that $a_{jt} = 0$ for any establishment at the industry-year mean.

evident, has deteriorated. The growth differential between high- and average-productivity firms has fallen from 25 percentage points (1996-1999 average) to below 21 percentage points (2011-2013), while the exit probability differential has gone from 7.6 to 6.7 percentage points.

An alternative way to quantify the substantive significance of our estimated decline in responsiveness is to consider what it implies for aggregate job reallocation. The estimated (empirical) responsiveness coefficient for TFPS reported on Table 1 declines by roughly 30 percent from the 1980s to the 2000s. There is also an accompanying increase in the standard deviation of TFPS from 0.46 to 0.51. Combining these effects implies a decline in the growth differential between high-productivity (i.e., one standard deviation above the mean) establishments and average establishments of about 24 percent (holding all else equal).⁴⁰ This suggests that declining responsiveness can account for a decline in employment growth rate dispersion of about 24 percent. Decker et al. (2016b) find that time series of employment growth rate dispersion closely tracks the time series of job reallocation rates (correlation of 0.92), and changes in growth rate dispersion are approximately proportional to changes in reallocation (see Figure A.1 from that paper). Since actual job reallocation declined in manufacturing by somewhat less than 20 percent over this time (see Figure 1), declining responsiveness is roughly sufficient to explain the overall decline in job reallocation. Two other exercises provide more confidence that this back-of-the-envelope calculation is reasonable. First, in an unreported exercise, we conduct a diff-in-diff counterfactual for aggregate job reallocation, inspired by our productivity diff-in-diff described above, and find broadly similar results.⁴¹ Second, we construct an alternative estimate of this effect using the calibrated model (see Appendix I.D) and find similar results.

Table 1 and Figure 4 strongly demonstrate that responsiveness has weakened among U.S. businesses, and the exercises just described suggest that this decline in responsiveness is

⁴⁰ As noted above, since establishment TFPS is deviated from the industry mean, we can obtain the relative employment growth of an establishment that is one standard deviation above its industry's mean TFPS versus an establishment at the mean by simply multiplying the growth responsiveness coefficient by the standard deviation of TFPS. Using estimates from decade dummies on Table 1, this differential is $0.46 * 0.29 = 0.1334$ in the 1980s and $0.51 * 0.20 = 0.102$ in the 2000s; 0.102 is a 24 percent decline from 0.1421. Note that this exercise differs from the previous one described on Figure 4 in that we now allow TFPS dispersion to vary over time.

⁴¹ For this counterfactual, we compute $\Delta_t^{t+1}(JR) = \sum_j \hat{\theta}_{jt+1}^{80s} \hat{g}_{it+1}^{80s} - \sum_j \hat{\theta}_{jt+1}^{00s} \hat{g}_{it+1}^{00s}$ where the predicted employment shares $\hat{\theta}_{jt+1}$ and growth rates \hat{g}_{jt+1} are based on the actual distribution of productivity and employment in t using the estimated coefficients from the model with decade dummies. That is, we generate a series yielding the change in the distribution of employment growth rates predicted by our estimated responsiveness models; this counterfactual series implies a decline in job reallocation similar to the aggregate decline observed on Figure 1.

economically significant. We observe weakening responsiveness to four independent measures of establishment- or firm-level productivity, and we see the decline in three different ways: negative linear trend estimates, a negative and significant step down in the second half of the sample versus the first half (using the post-1997 indicator), and significantly different responsiveness coefficients in the 1980s, the 1990s, and the 2000-onward period. Overall employment growth responsiveness has weakened, as has the sensitivity of establishment or firm exit.

As discussed above, in an environment characterized by adjustment costs, there should also be a decline in responsiveness to the *innovation* to, or change in, productivity. As previously noted, our data are not ideally suited to measuring productivity changes; our ASM-CM sample is representative in any specific year but is not designed to be longitudinally representative. Nevertheless, to further explore robustness of our main results, we estimate our manufacturing regressions replacing the level of productivity (a_{jt-1}) with the innovation to productivity (given by $\eta_{jt-1} = a_{jt-1} - \rho_a a_{jt-2}$). We estimate innovation regressions in our empirical dataset and report the results on Table 2, focusing on our preferred productivity measure TFPS.⁴² Note that this exercise significantly reduces our sample size (from over 2 million establishment-year observations to fewer than 1 million). Regardless, though, we still observe a decline in responsiveness of employment growth to TFPS innovations, with a particular step down between the 1990s and the 2000s. In other words, our main results are broadly robust to the use of productivity innovations rather than productivity levels. Moreover, in unreported exercises we find similar results in a regression using first differences in TFPS rather than innovations and obtain very similar results.

Taken together with the evidence of rising productivity dispersion—in multiple productivity measures—discussed above and viewed through the lens of canonical models of business dynamics, these results suggest that the costs or incentives to adjust employment in response to changing economic circumstances have changed over time. More broadly, the decline in the aggregate pace of job reallocation—the motivating fact for this paper—does not

⁴² Table 1 shows that the various TFP measures deliver broadly similar results, so in all remaining empirical exercises we dispense with our TFPP and TFPR measures and report only TFPS results. As noted above, TFPS—the TFP measure based on elasticities from factor shares of revenue—is our preferred TFP measure, as it allows for endogenous prices while avoiding the imprecision of revenue function estimation. We also continued to report specifications based on RLP to gain perspective on non-manufacturing activity.

appear to be a result of a less turbulent idiosyncratic environment faced by businesses; rather, businesses have become more sluggish in responding to shocks.

However, as noted above, one critical change in the composition of U.S. businesses in recent decades may be affecting these results: the secular decline in young firm activity. If young firms are typically more responsive to shocks than are more mature firms, overall responsiveness would decline as young firm activity falls. The potential for firm age-based composition effects to affect our results is possibly a significant limitation of the exercises presented on Table 1. It is of interest, therefore, to study changing responsiveness within firm age groups, which we do next.

C. “Responsiveness” hypothesis: Young versus mature firms

We now expand the regression model given by equation (10) to allow differential responsiveness—and responsiveness trends—between young and mature firms, where we define “young” firms as those with age less than five:

$$g_{jt+1} = (\beta_1^y a_{jt} + T^y(a_{jt}, t) + \beta_2^y e_{jt} + T^y(e_{jt}, t))\mathbb{I}_{\{y=1\}} + (\beta_1^m a_{jt} + T^m(a_{jt}, t) + \beta_2^m e_{jt} + T^m(e_{jt}, t))\mathbb{I}_{\{m=1\}} + X'_{jt}\Theta + \varepsilon_{jt+1}, \quad (12)$$

where y indicates young firms (those with age less than five), m indicates mature firms (those with age five or greater), and each $\mathbb{I}_{\{\cdot\}}$ is a corresponding age dummy indicator. Note that this indicator refers to *firm* age, even in establishment-level (manufacturing) regressions; given firm information from the LBD, we can identify the age of the firm to which any given establishment belongs. $T^y(a_{jt}, t)$ and $T^m(a_{jt}, t)$ (and the corresponding effects for e_{jt}) are defined as in (11) with the addition of firm age superscripts on all relevant coefficients. We also include interactions of the cyclical controls with firm age in X'_{jt} .⁴³

Equation (12) allows us to study overall responsiveness and how it has changed over time *within* firm age groups, thereby controlling for changes in the firm age composition of U.S. businesses. Decker et al. (2014) study firm lifecycle dynamics and find that the critical characteristics that are unique to young firms—growth skewness and dispersion—are most pronounced prior to age five, with the distribution of growth settling somewhat thereafter. Based

⁴³ We use the same revenue elasticities at the industry level for measuring TFPS for both young and mature firms. Foster, Haltiwanger, and Syverson (2016) present evidence suggesting markups vary over the life cycle with smaller markups for younger firms. We are neglecting any such effects in our analysis.

on this, we use two firm age groups: establishments owned by firms younger than five, and those owned by firms aged five and above.⁴⁴

Table 3 reports results from the regression in (12) for TFPS and RLP. As before, we report results with employment growth as the dependent variable (Panel A, on the left) and exit as the dependent variable (Panel B, on the right). For TFPS regressions in manufacturing, we report both the linear trend coefficient and the decade dummy specification (with decade-equality t test p values). As can be seen in all specifications—TFPS and RLP, with both growth and exit as dependent variables—young firms are indeed more responsive than mature firms (even within decades). In other words, young firms face more intense selection. As such, some portion of the decline in responsiveness reported on Table 1 does indeed reflect the changing age composition of firms. However, as the trend and decade coefficients demonstrate, responsiveness has declined over time within firm age groups.

Responsiveness has particularly declined among young firms, which have historically been more responsive. The 2000s growth coefficient for young firms, 0.25, is weaker than the initial 1980s coefficient for mature firms, 0.27. Following the exercise used for Figure 4, the growth differential for young firms with TFPS one standard deviation above their industry-year mean has declined from over 17 percentage points in the 1980s to just over 12 percentage points in the 2000s, while the exit risk differential has narrowed from 4.9 to 3.3 percentage points.

Selection—the relationship between productivity and business growth and survival—is critical to aggregate productivity growth, and selection has historically been more intense for young firms. The significant weakening of selection among young firms is therefore particularly concerning. Before quantifying the implications of our results for aggregate productivity, however, we will first explore patterns of responsiveness for high-tech industries.

D. “Responsiveness” hypothesis: High-tech

As discussed in the introduction and shown on Figure 1, patterns of reallocation in the high-tech sector have differed from the broader economy in recent decades. In particular, in high-tech the pace of reallocation rose during the 1980s and 1990s before declining in the 2000s. Given our shocks vs. responsiveness framework, the reallocation patterns lead us to expect productivity responsiveness to behave similarly; that is, we expect productivity responsiveness in the high-tech sector to strengthen during the 1980s and 1990s, then weaken thereafter.

⁴⁴ Disclosure limitations also preclude a more detailed age analysis.

We now estimate equation (12) separately for high-tech and non-tech businesses (see the data discussion for details on industry classification). We consider these exercises to be our richest specification, as they fully exploit the data along dimensions that are of crucial importance for reallocation and productivity patterns. Again, we report results using the TFPS and RLP productivity concepts.

Table 4 reports the results of these regressions, where we report only growth regressions (omitting exit regressions for brevity; recall that our DHS growth variable is inclusive of exit). We focus first on TFPS results, the first four columns of the table. While the results for non-tech establishments (the first two columns) are similar to those of the economy generally (shown on Table 3), responsiveness patterns in high-tech (the third and fourth columns) are different, in a manner consistent with aggregate reallocation patterns. This can clearly be seen in the decade-specific responsiveness coefficients: responsiveness rises between the 1980s and 1990s and steps back down in the 2000s. This rising and falling pattern is particularly evident among young high-tech firms. Figures 5a and 5b report growth differentials with the method from Figure 4; among young high-tech firms in manufacturing, the employment growth differential between high-productivity establishments and average establishments rose from 12 percentage points in the 1980s to over 15 percentage points in the 1990s then fell to less than 8 percentage points in the 2000s.⁴⁵

The last two columns of Table 4 report regressions using RLP as the productivity concept. Our RLP data only begin in 1996, but the high-tech responsiveness pattern is evident even in the linear trend coefficients (δ^y and δ^m), particularly among young firms. Figures 5c and 5d report the coefficients in terms of growth rate differentials, averaged for 1996-1999 and 2011-2013; among young high-tech firms, the growth rate differential between high-productivity firms and their industry average has declined from 30 to 23 percentage points.

Table 4 and Figure 5 tell a rich story about productivity responsiveness and selection in recent decades. Consistent with patterns of job reallocation, responsiveness among non-tech businesses has declined steadily and significantly in recent decades, particularly among young firms, which have historically faced intense selection but increasingly behave more like mature firms. In high-tech, we observe rising responsiveness from the 1980s into the 1990s then falling

⁴⁵ Recall that this exercise compares high-productivity establishments—those with TFPS that is one standard deviation above their industry-year mean—to establishments at the industry-year mean.

responsiveness thereafter. These patterns are consistent with aggregate patterns of job reallocation. More broadly, Table 4 is consistent with the results of Tables 1-3, which tell a story of pervasive decline in productivity responsiveness by the 2000s.

Strikingly, the high-tech productivity responsiveness pattern is also consistent with patterns of aggregate productivity growth during the 1990s and the 2000s, as can be seen in the decade coefficients for high-tech manufacturing in Table 4. As shown by Fernald (2014), aggregate productivity growth in the U.S. increased during the 1990s before stepping down in the early-to-mid 2000s, driven largely by industries that produce or heavily use ICT products.

V. Implications for aggregate productivity

The rise and fall in U.S. productivity growth in recent decades is the subject of a large literature. Using simple accounting decompositions in the RE-LBD, Decker et al. (2017) find some evidence that the post-2000 productivity slowdown is related in part to slowing job reallocation, which has apparently reduced productivity gains from allocative efficiency improvements.⁴⁶ We now study the relationship between job reallocation and aggregate productivity more directly using our estimated responsiveness coefficients. The results documented above imply that productivity selection has weakened, and the pace of the flow of resources from less-productive businesses to more-productive businesses has declined. In Section II, we use our model to illustrate a method for quantifying the aggregate productivity implications of changing responsiveness. In this section, we apply that method to our empirical estimates (see Section II for details about the method).

We first focus on our manufacturing sample. When we estimated equation (12) we quantified changes in responsiveness between the 1980s, the 1990s, and the 2000s. Those results are reported on Table 4 for both high-tech and non-tech industries. Using those estimates and following equations (3) through (5), we construct a diff-in-diff measure of the effect of changing responsiveness between the 1980s and the 1990s:

$$\Delta_t^{t+1}(1990s) = (\hat{a}_{t+1}^{90s} - a_t) - (\hat{a}_{t+1}^{80s} - a_t) = \sum_j \hat{\theta}_{jt+1}^{90s} a_{jt} - \sum_j \hat{\theta}_{jt+1}^{80s} a_{jt}, \quad (13)$$

where $\hat{\theta}_{jt+1}^{80s}$ is the employment share of establishment j in year $t + 1$ as predicted by equation (12) assuming the 1980s rate of responsiveness, and $\hat{\theta}_{jt+1}^{90s}$ is the employment share of

⁴⁶ Decker et al. (2017) use the dynamic Olley-Pakes decomposition of Melitz and Polanec (2015) and the Foster, Haltiwanger, and Krizan (2001) decomposition.

establishment j in year $t + 1$ as predicted by equation (12) assuming the 1990s rate of responsiveness. That is, for every establishment in the data, we use *observed* values of all variables necessary for equation (12), including (log) initial productivity a_{jt} , (log) initial employment e_{jt} , and other establishment information including firm age and industry; taken together, these variables predict an employment growth rate \hat{g}_{jt+1} which, when multiplied by initial employment, implies an employment share $\hat{\theta}_{jt+1}$ for each establishment j .⁴⁷ We also construct a diff-in-diff measure comparing the 2000s and the 1980s:

$$\Delta_t^{t+1}(2000s) = (\hat{a}_{t+1}^{00s} - a_t) - (\hat{a}_{t+1}^{80s} - a_t) = \sum_j \hat{\theta}_{jt+1}^{00s} a_{jt} - \sum_j \hat{\theta}_{jt+1}^{80s} a_{jt}, \quad (14)$$

Figure 6a reports the counterfactual results ($\Delta_t^{t+1}(1990s)$ and $\Delta_t^{t+1}(2000s)$) for all *establishments* together and for non-tech and high-tech establishments separately, using our preferred TFPS measure of TFP. The first bar shows that the change in aggregate productivity between years t and $t + 1$ is more than one log point lower under the 1990s rate of productivity responsiveness than it is under the 1980s rate of responsiveness. This 1-percentage-point effect indicates how much higher average aggregate productivity in the 1990s would have been (in each year $t + 1$ relative to the prior year t) if responsiveness were immediately to return to its 1980s strength, given the actual distribution of productivity we observe in each initial year t . The second bar shows that by the 2000s, responsiveness has fallen so much that a return to 1980s responsiveness would increase productivity in each year relative to the prior year by more than 2 log points.

The next two bars on Figure 6a report results for the high-tech sector. Recall from Table 4 that responsiveness strengthened in the high-tech sector from the 1980s to the 1990s, which is reflected in the counterfactual. Relative to the 1980s, increased responsiveness in the 1990s adds about 1 log point to productivity in each year (relative to the prior year) in the high-tech sector. However, by the 2000s, responsiveness has declined such that a return to 1980s responsiveness would boost aggregate productivity in each year relative to the prior year by nearly 2 log points. The counterfactual pattern seen in high-tech mimics the pattern of aggregate productivity growth

⁴⁷ We set all cyclical effects to zero, as if predicting growth under neutral business cycle conditions. Foster, Grim, and Haltiwanger (2016) show that productivity responsiveness is cyclical; our intention is to conduct a sort of steady state analysis, avoiding the influence of cyclical patterns. We also permit the impact of lagged employment to vary over time in generating these predicted shares while we keep the responsiveness to productivity constant at the 1980s level. Results are similar if the counterfactual is conducted with the coefficients on lag employment fixed at their 1980s values.

for the U.S. broadly during the 1990s and 2000s; that well-documented acceleration then deceleration was driven primarily by ICT-producing and using industries (Fernald (2014)). Our results suggest that changing productivity responsiveness may be one part of the explanation for these recent U.S. productivity dynamics. More specifically, though, Figure 6a illustrates that the changes in responsiveness we document have likely had significant implications for aggregate productivity dynamics in the U.S.

Finally, we construct similar counterfactuals using our economywide data. We did not estimate decade-specific responsiveness coefficients in the economywide sample; rather, we relied upon the simple linear trend coefficient. Those results are described in the “RLP” columns of Table 4. In principle, using the trend coefficient on responsiveness, we could construct a counterfactual for every year in our data, 1996-2013. For brevity, we instead average the trend-implied responsiveness coefficients for multi-year periods and report the diff-in-diff for the first and last of these periods. In each period, we compare responsiveness to its initial level (i.e., $Trend_t = 0$ in equation (11)). We report only economywide figures, not breaking out high-tech industries, since the time window is too short to interpret differences.

Figure 6b reports these economywide diff-in-diff objects in the initial multi-year period (1996-1999) and the final period (2011-2013). The effect of declining responsiveness grows to more than 3 log points by 2013. This suggests that by 2013, if responsiveness immediately returned to its 1996 strength, aggregate productivity relative to 2012 would have been 3 log points higher. This is a much stronger number than what we observe in manufacturing. Recall, however, that we take as given the existing revenue productivity distribution in any given year; for example, the diff-in-diff for year 2013 takes as given the actual distribution of revenue per worker in 2012, which is in fact endogenous to changing responsiveness over time. We therefore urge caution in the interpretation of these results; however, they demonstrate the quantitative importance of changing responsiveness.

VI. Robustness exercises

Tables 1, 2, 3, and 4 demonstrate robustness of our responsiveness results to alternative productivity measures, including innovations to productivity. We show that our changing responsiveness evidence holds up within firm age groups and is evident in a variety of time trend specifications. Our results typically have extremely high statistical significance. This robustness

notwithstanding, we now briefly explore two other questions: responsiveness of other factors, and robustness to employment measurement issues.

A. *Investment responsiveness*

One possible explanation for declining employment responsiveness is that business models have changed such that businesses increasingly respond to idiosyncratic shocks by adjusting factors other than employment. This may be thought of as a capital/labor substitution mechanism. While our economywide firm dataset lacks information on factors other than labor, the manufacturing data are much richer. Here we focus on manufacturing establishments and study the responsiveness of equipment investment to productivity over time. We estimate specification (12) replacing the DHS employment growth rate with the investment rate I_{jt}/K_{jt} , where I_{jt} is equipment investment throughout year t , and K_{jt} is the stock of capital equipment at the beginning of the year. Models of business dynamics with adjustment costs produce policy functions and intuition for capital investment that are similar to the policy functions and intuition for employment growth.⁴⁸

We estimate the investment responsiveness specification separately for high-tech and non-tech establishments, and for brevity we report only the regression in which the time functions $T^y(a_{jt}, t)$ and $T^m(a_{jt}, t)$ are specified as decade-specific coefficients (the last element in (11)). Table 5 reports the results based again on our preferred TFP measure, TFPS, standing in for a_{jt} . In the first column, among all industries we see rising investment responsiveness from the 1980s to the 1990s followed by a drop in the 2000s. This pattern is consistent with capital/labor substitution between the 1980s and 1990s while employment responsiveness was falling, after which investment joins employment in becoming less responsive in the 2000s. The non-tech column is similar.

The high-tech column is striking: while responsiveness among high-tech establishments was stronger than the rest of the manufacturing sector during the 1980s, by the 2000s it is far weaker than the rest of manufacturing and is in fact no longer statistically significant (though the difference between the 1990s and the 2000s is significant). We observe a similar pattern among

⁴⁸ To make results comparable and minimize disclosure issues we use exactly the same specification as (12) just replacing the dependent variable. Unreported specifications where we also control for the initial capital stock each period (allowing such effects to vary over time as with other variables) produce similar results. Unreported results also show that adding the initial capital stock variable in a similar way to the employment growth regressions in Table 4 are very similar to those reported in Table 4.

both young and mature firms. In other words, productivity selection for investment appears completely absent in high-tech manufacturing during the 2000s. Similarly to our employment results, we can quantify investment responsiveness in terms of the differential between high-productivity establishments (those that are one standard deviation more productive than their industry mean) and average (within industry) establishments. Among high-tech young firm establishments, this differential in investment rates rose from 4.8 percentage points to 7.4 percentage points from the 1980s to the 1990s before falling close to zero in the 2000s.

In recent years researchers have given increasing focus to intangible capital (e.g., Corrado, Hulten, and Sichel (2009), Haskel and Westlake (2017)). It is possible that weakening responsiveness of employment and equipment investment to productivity has been accompanied by changes in the responsiveness of intangible capital investment; however, we cannot answer this question with our data. The results for equipment investment suggest that factor demand responsiveness generally may have changed in a pattern consistent with the time series of job reallocation rates.

B. Employment measurement error and division bias

Employment appears in several places in our responsiveness regressions, including both the left-hand side and the right-hand side. The prolific inclusion of initial employment raises questions about division bias in the presence of measurement error, which could introduce bias in our estimated responsiveness coefficients.

We first note that our analysis of *exit* uses a dependent variable equal to one if a plant or firm exits between t and $t + 1$ based on productivity and employment measures in t . Measurement error in productivity and employment in t will generate attenuation bias but not division bias in this specification. Our results on exit (the extensive margin of selection) are consistent with our analysis that incorporates both extensive and intensive margins (i.e., when we use DHS growth rates as the dependent variable).

For the specifications with the DHS growth rate as the dependent variable, we explore the potential consequences of measurement error in employment in period t by using lagged employment (i.e., $t - 1$) as an instrument (where the E_{t-1} is measured from the LBD). This potentially mitigates measurement error division bias problems since E_{t-1} does not appear in the

DHS growth rate measuring growth from t to $t + 1$.⁴⁹ The potential division bias carries over to the specifications with RLP where the denominator of RLP in period t is E_t from the LBD. (Recall, however, that the employment values used for constructing TFPS, TFPP, and TFPR are total hours from the ASM/CM, a different source of employment from the LBD). For the RLP regressions, we also instrument RLP in period t with lagged RLP in period $t - 1$. We report these results on Table A2 of Appendix III. The IV specifications show results that are broadly similar to our main regression results.

Additionally, in Appendix I, we describe several model experiments in which the econometrician (but not firm managers) observes employment with some measurement error. Measurement error does introduce biases into our responsiveness coefficient, with the direction depending on whether there is error in both employment and productivity or just productivity. Changes in measurement error over time could therefore be responsible for the decline in responsiveness we document in our main empirical exercises, though our model results suggest that the measurement error, and changes in it, would have to be extremely large relative to what we view as plausible. We cannot test for this possibility. However, we have no reason to expect measurement error to increase or decrease over time in a way that would align with our findings; and the measurement error would have to behave differently in high-tech and non-tech businesses in a way that perfectly corroborates the patterns of aggregate job reallocation we observe. We find this unlikely and prefer to interpret our empirical results at face value.

VII. Conclusion

Resource reallocation plays a critical role in productivity dynamics. The U.S. has seen a decline in the pace of job reallocation in recent decades that has proven difficult to understand. We study changing patterns of reallocation by drawing insight from canonical models of firm dynamics. In such models, a decline in the pace of job reallocation can arise from one of two sources: (1) the dispersion or volatility of idiosyncratic *shocks* faced by businesses—the business-specific conditions that drive hiring and downsizing decisions—could have declined; or (2) business-level *responsiveness* to idiosyncratic conditions may have weakened.

We show that shock dispersion has not declined but has actually risen, and business-level responsiveness to shocks—in terms of employment growth and survival—has weakened. Our

⁴⁹ Recall that growth dependent variable is $g_{t+1} = (E_{t+1} - E_t) / (.5E_{t+1} + .5E_t)$.

finding of weakening responsiveness is robust to alternative productivity measures, three different time-trend specifications, and changing firm age composition. It holds for both the level of shocks and the innovations to shocks. Equipment investment has also become less responsive, at least in the last two decades, suggesting that capital/labor substitution is not a likely explanation, though we cannot rule out substitution into intangible production factors.

Weakening responsiveness to productivity has significant implications for aggregate productivity since it reflects a slower pace of reallocation in response to idiosyncratic productivity shocks. That is, productivity selection has weakened. We quantify these effects using a diff-in-diff approach that isolates the effect of weaker responsiveness on aggregate productivity, finding that the effect is substantial and amounts to a drag of about 2 log points each year in the last decade in manufacturing and a similar order of magnitude in the economy broadly. Our productivity results strongly suggest that the decline in job reallocation has not been benign for living standards.

We view our results on declining responsiveness as independently significant and as substantial progress on questions about changing business dynamics in the U.S. Discovering the causes of declining responsiveness is important but beyond the scope of this paper. Our model framework suggests that rising factor adjustment costs may play a role, since rising costs imply both weaker responsiveness and higher dispersion of revenue-based productivity measures (which we observe). This could be the result of regulatory changes that affect the cost of hiring or downsizing or of other changes in the economic environment.

As noted previously, however, adjustment costs are not the only possible explanation. In particular, any changes in the wedges that impede the equalization of marginal revenue products may be at work. A rise in adjustment costs is one potential candidate for rising wedges. Alternatively, increasing market power and rising markups (as studied by De Loecker, Eeckhout, and Unger (2018)) could drive weaker responsiveness (see Appendix I). More broadly, any changes in the costs of or incentives for hiring and downsizing could affect responsiveness. We view identifying the source of the rising wedges as a critical avenue for future research.

References

- 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*.
- 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.
- Barro, Robert J. 1972. "A Theory of Monopolistic Price Adjustment." *The Review of Economic Studies* 39(1): 17-26.
- Barth, Erling, Alex Bryson, James C. Davis, and Richard Freeman. 2016. "It's Where You Work: Increases in the Dispersion of Earnings Across Establishments and Individuals in the United States." *Journal of Labor Economics* 34 no. 2 pt. 2: S67-S97.
- Berger, David and Joseph Vavra. Forthcoming. "Shocks vs. Responsiveness: What Drives Time-Varying Dispersion," *Journal of Political Economy*.
- Bils, Mark, Peter J. Klenow, and Cian Ruane. 2017. "Misallocation or Mismeasurement?" Working paper.
- Blackwood, Glenn, Lucia Foster, Cheryl Grim, John Haltiwanger and Zoltan Wolf. 2019. "Macro and Micro Dynamics of Productivity: From Devilish Details to Insights", Working Paper.
- Bloom, Nicholas. 2009. "The Impact of Uncertainty Shocks." *Econometrica* 77 no. 3:623-85.
- Bloom, Nicholas, Max Floetotto, Nir Jaimovich, Itay Saporta-Eksten, and Stephen J. Terry. 2018. "Really Uncertain Business Cycles." *Econometrica* 86 no. 3:1031-65.
- 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.
- 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 no. 1:37-50.
- Corrado, Carol, Charles Hulten, and Daniel Sichel. 2009. "Intangible Capital and Economic Growth." *Review of Income and Wealth* 55 no. 3:661-685.
- 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, Jan Eeckhout and Gabriel Unger. 2018. "The Rise of Market Power and the Macroeconomic Implications." Working paper.
- 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. "Declining Business Dynamism: Implications for Productivity?" Hutchins Center Working Paper #23.
- 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).
- Decker, Ryan A., John Haltiwanger, Ron S. Jarmin, and Javier Miranda. 2018. "Changing Business Dynamism and Productivity: Shocks vs. Responsiveness." NBER Working Paper no. 24236.

- Elsby, Michael W. L. and Ryan Michaels. 2013. "Marginal Jobs, Heterogeneous Firms, and Unemployment Flows." *American Economic Journal: Macroeconomics* 5 no. 1:1-48.
- 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* 34 no. S1 (Part 2, January 2016):S293-S331.
- 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*, January, 91-129.
- Gopinath, Gita, Sebnem Kalemli-Ozcan, Loukas Karabarbounis and Carolina Villegas-Sanchez 2017. "Capital Allocation and Productivity in South Europe." *Quarterly Journal of Economics* 132 no. 4: 1915-1967.
- 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, 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. 2017. "High Growth Young Firms: Contribution to Job Growth, Revenue Growth and Productivity" 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, Ron S. Jarmin, and Javier Miranda. 2013. "Who Creates Jobs? Small versus Large versus Young." *The Review of Economics and Statistics* 95 no. 2: 347-361.
- Haskel, Jonathan and Stian Westlake. 2017. *Capitalism without Capital: The Rise of the Intangible Economy*. Princeton University Press.
- 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.
- Hyatt, Henry R. and James Spletzer. 2013. "The Recent Decline in Employment Dynamics." *IZA Journal of Labor Economics* 2(3): 1-21.
- Iltut, 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., 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. 2018. "Demographic Origins of the Startup Deficit." Mimeo.
- 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.
- 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.

Table 1: Business-level employment growth responsiveness has weakened

<i>Panel A – dependent variable: employment growth (g_{jt+1}) from t to $t+1$</i>										
	TFPS (revenue share based)			TFPP (proxy method)			TFPR (cost share based)			RLP (revenue per worker)
Productivity: β_1	0.2965 (0.0097)	0.2905 (0.0068)		0.2086 (0.0086)	0.1876 (0.0060)		0.2040 (0.0094)	0.1779 (0.0067)		0.2762 (0.0003)
Prod*trend: δ	-0.0035 (0.0005)			-0.0043 (0.0004)			-0.0046 (0.0005)			-0.0029 (0.0000)
Prod*post-97: γ_{97}		-0.0952 (0.0084)			-0.0958 (0.0074)			-0.0981 (0.0081)		
Prod*1980s: λ_{80s}			0.2859 (0.0095)			0.2185 (0.0086)			0.1939 (0.0094)	
Prod*1990s: λ_{90s}			0.2462 (0.0060)			0.1053 (0.0052)			0.1212 (0.0058)	
Prod*2000s: λ_{00s}			0.2001 (0.0059)			0.0995 (0.0051)			0.0820 (0.0054)	
p value: $\lambda_{80s} = \lambda_{90s}$			0.00			0.00			0.00	
p value: $\lambda_{80s} = \lambda_{00s}$			0.00			0.00			0.00	
p value: $\lambda_{90s} = \lambda_{00s}$			0.00			0.43			0.00	
Obs. (thousands)	2,375	2,375	2,375	2,375	2,375	2,375	2,375	2,375	2,375	58,700

<i>Panel B – dependent variable: exit between t and $t+1$</i>										
	TFPS (revenue share based)			TFPP (proxy method)			TFPR (cost share based)			RLP (revenue per worker)
Productivity: β_1	-0.0757 (0.0009)	-0.0801 (0.0030)		-0.0830 (0.0038)	-0.0781 (0.0027)		-0.0721 (0.0042)	-0.0664 (0.0030)		-0.0857 (0.0001)
Prod*trend: δ	0.0009 (0.0002)			0.0014 (0.0002)			0.0014			0.0007 (0.0000)
Prod*post-97: γ_{97}		0.0340 (0.0037)			0.0352 (0.0033)			0.0330 (0.0036)		
Prod*1980s: λ_{80s}			-0.0773 (0.0042)			-0.0868 (0.0038)			-0.0714 (0.0042)	
Prod*1990s: λ_{90s}			-0.0586 (0.0026)			-0.0473 (0.0022)			-0.0430 (0.0025)	
Prod*2000s: λ_{00s}			-0.0517 (0.0026)			-0.0478 (0.0023)			-0.0370 (0.0024)	
p value: $\lambda_{80s} = \lambda_{90s}$			0.00			0.00			0.00	
p value: $\lambda_{80s} = \lambda_{00s}$			0.00			0.00			0.00	
p value: $\lambda_{90s} = \lambda_{00s}$			0.06			0.89			0.09	
Obs. (thousands)	2,375	2,375	2,375	2,375	2,375	2,375	2,375	2,375	2,375	58,700

Note: All coefficients statistically significant with $p < 0.01$. TFPS, TFPP, and TFPR columns are establishment regressions in manufacturing for 1981-2013. RLP columns are economywide firm regressions for 1997-2013. All regressions include controls described in equation (10) and related text.

Table 2: Employment growth has also become less responsive to productivity innovations

	TFPS η_{jt}
Prod*1980s: λ_{80s}	0.3970 (0.0279)
Prod*1990s: λ_{90s}	0.3909 (0.0124)
Prod*2000s: λ_{00s}	0.2999 (0.0126)
p value: $\lambda_{80s} = \lambda_{90s}$	0.84
p value: $\lambda_{80s} = \lambda_{00s}$	0.00
p value: $\lambda_{90s} = \lambda_{00s}$	0.00
Observations (thousands)	854

Note: Regression of employment growth on TFPS innovation and controls as described in equation (10) and related text. All coefficients are statistically significant with $p < 0.01$.

Table 3: Employment growth responsiveness has weakened within firm age groups

	<i>Panel A – dependent variable: emp. growth (g_{jt+1}) from t to $t+1$</i>		<i>Panel B – dependent variable: exit between t and $t+1$</i>		
	TFPS	RLP	TFPS	RLP	
Prod*Young: β_1^y	0.4069 (0.0137)	0.3217 (0.0003)	-0.1075 (0.0060)	-0.1065 (0.0001)	
Prod*Young*trend: δ^y	-0.0054 (0.0006)	-0.0034 (0.0000)	0.0014 (0.0003)	0.0009 (0.0000)	
Prod*Mature: β_1^m	0.2722 (0.0096)	0.2493 (0.0003)	-0.0690 (0.0042)	-0.0747 (0.0001)	
Prod*Mature*trend: δ^m	-0.0029 (0.0005)	-0.0024 (0.0000)	0.0007 (0.0002)	0.0005 (0.0000)	
Prod*Young*1980s: λ_{80s}^y		0.3666 (0.0136)		-0.1020 (0.0059)	
Prod*Young*1990s: λ_{90s}^y		0.3603 (0.0092)		-0.0898 (0.0039)	
Prod*Young*2000s: λ_{00s}^y		0.2542 (0.0093)		-0.0689 (0.0039)	
Prod*Mature*1980s: λ_{80s}^m		0.2710 (0.0094)		-0.0727 (0.0042)	
Prod*Mature*1990s: λ_{90s}^m		0.2185 (0.0058)		-0.0529 (0.0025)	
Prod*Mature*2000s: λ_{00s}^m		0.1941 (0.0059)		-0.0507 (0.0026)	
p value: $\lambda_{80s}^y = \lambda_{90s}^y$ (young)		0.67		0.06	
p value: $\lambda_{80s}^y = \lambda_{00s}^y$ (young)		0.00		0.00	
p value: $\lambda_{90s}^y = \lambda_{00s}^y$ (young)		0.00		0.00	
p value: $\lambda_{80s}^m = \lambda_{90s}^m$ (mature)		0.00		0.00	
p value: $\lambda_{80s}^m = \lambda_{00s}^m$ (mature)		0.00		0.00	
p value: $\lambda_{90s}^m = \lambda_{00s}^m$ (mature)		0.00		0.55	
Observations (thousands)	2,375	2,375	2,375	2,375	58,700

Note: All coefficients statistically significant with $p < 0.01$. TFPS columns are establishment-level regressions in manufacturing data for 1981-2013. RLP columns are firm-level regressions on economywide data for 1997-2013. All regressions include controls described in equation (10) and related text. Young firms have age less than five.

Table 4: Responsiveness patterns differ between high-tech and non-tech industries

	<i>Dependent variable: employment growth (g_{jt+1}) from t to $t+1$</i>			
	TFPS		RLP	
	Non-tech	High-tech	Non-tech	High-tech
Prod*Young: β_1^y	0.4171 (0.0145)	0.3195 (0.0387)	0.3207 (0.0003)	0.3436 (0.0015)
Prod*Young*trend: δ^y	-0.0054 (0.0007)	-0.0058 (0.0020)	-0.0033 (0.0000)	-0.0054 (0.0001)
Prod*Mature: β_1^m	0.2786 (0.0101)	0.1969 (0.0302)	0.2481 (0.0003)	0.2813 (0.0014)
Prod*Mature*trend: δ^m	-0.0030 (0.0005)	-0.0022† (0.0016)	-0.0023 (0.0000)	-0.0044 (0.0001)
Prod*Young*1980s: λ_{80s}^y	0.3801 (0.0144)	0.2488 (0.0371)		
Prod*Young*1990s: λ_{90s}^y	0.3637 (0.0097)	0.3190 (0.0331)		
Prod*Young*2000s: λ_{00s}^y	0.2637 (0.0095)	0.1575 (0.0332)		
Prod*Mature*1980s: λ_{80s}^m	0.2792 (0.0099)	0.1657 (0.0293)		
Prod*Mature*1990s: λ_{90s}^m	0.2212 (0.0061)	0.1915 (0.0176)		
Prod*Mature*2000s: λ_{00s}^m	0.1964 (0.0060)	0.1340 (0.0216)		
p value: $\lambda_{80s}^y = \lambda_{90s}^y$ (young)	0.36	0.14		
p value: $\lambda_{80s}^y = \lambda_{00s}^y$ (young)	0.00	0.06		
p value: $\lambda_{90s}^y = \lambda_{00s}^y$ (young)	0.00	0.00		
p value: $\lambda_{80s}^m = \lambda_{90s}^m$ (mature)	0.00	0.45		
p value: $\lambda_{80s}^m = \lambda_{00s}^m$ (mature)	0.00	0.38		
p value: $\lambda_{90s}^m = \lambda_{00s}^m$ (mature)	0.00	0.04		
Observations (thousands)	2,239	136	55,870	2,826

Note: High-tech industries defined as in Hecker (2005). All coefficients statistically significant with $p < 0.01$ unless otherwise noted. TFPS columns are establishment-level regressions in manufacturing data for 1981-2013. RLP columns are firm-level regressions on economywide data for 1997-2013. All regressions include controls described in equation (10) and related text. Young firms have age less than five.

† Not statistically significant.

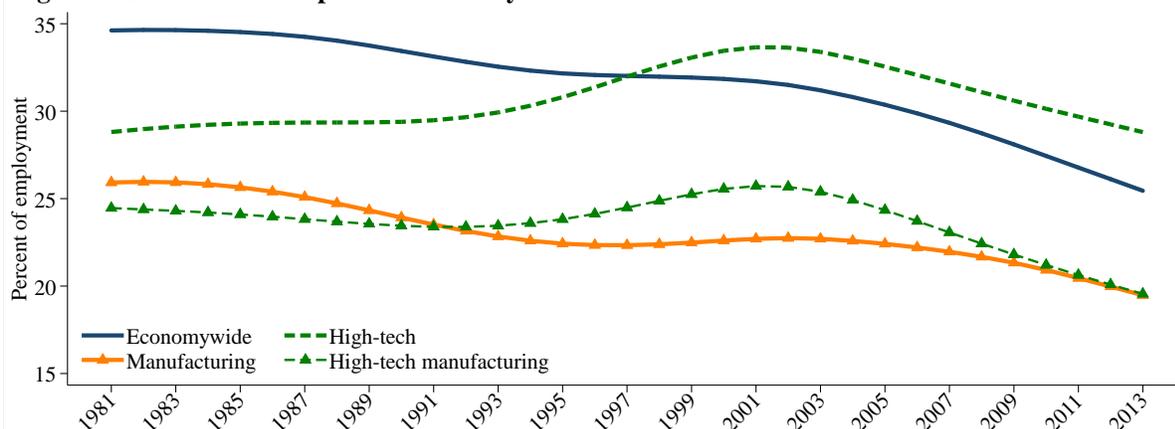
Table 5: Investment rate responsiveness has also weakened (manufacturing)

	All industries	Non-tech	High-tech
Prod*Young*1980s: λ_{80s}^y	0.0670 (0.0098)	0.0668 (0.0103)	0.1005 (0.0330)
Prod*Young*1990s: λ_{90s}^y	0.1768 (0.0139)	0.1785 (0.0147)	0.1550 (0.0312)
Prod*Young*2000s: λ_{00s}^y	0.1003 (0.0092)	0.1048 (0.0086)	0.0133† (0.0281)
Prod*Mature*1980s: λ_{80s}^m	0.0414 (0.0068)	0.0393 (0.0071)	0.0746 (0.0283)
Prod*Mature*1990s: λ_{90s}^m	0.1151 (0.0082)	0.1144 (0.0087)	0.1294 (0.0200)
Prod*Mature*2000s: λ_{00s}^m	0.0619 (0.0051)	0.0658 (0.0053)	-0.0020† (0.0194)
p value: $\lambda_{80s}^y = \lambda_{90s}^y$ (young)	0.00	0.00	0.22
p value: $\lambda_{80s}^y = \lambda_{00s}^y$ (young)	0.00	0.00	0.04
p value: $\lambda_{90s}^y = \lambda_{00s}^y$ (young)	0.00	0.00	0.00
p value: $\lambda_{80s}^m = \lambda_{90s}^m$ (mature)	0.00	0.00	0.12
p value: $\lambda_{80s}^m = \lambda_{00s}^m$ (mature)	0.02	0.00	0.03
p value: $\lambda_{90s}^m = \lambda_{00s}^m$ (mature)	0.00	0.00	0.00
Observations (thousands)	2,375	2,239	136

Note: Manufacturing only with TFPS productivity concept. High-tech industries defined as in Hecker (2005). All coefficients statistically significant with $p < 0.05$ (and almost always $p < 0.01$) unless otherwise noted. All regressions include controls described in equation (10) and related text.

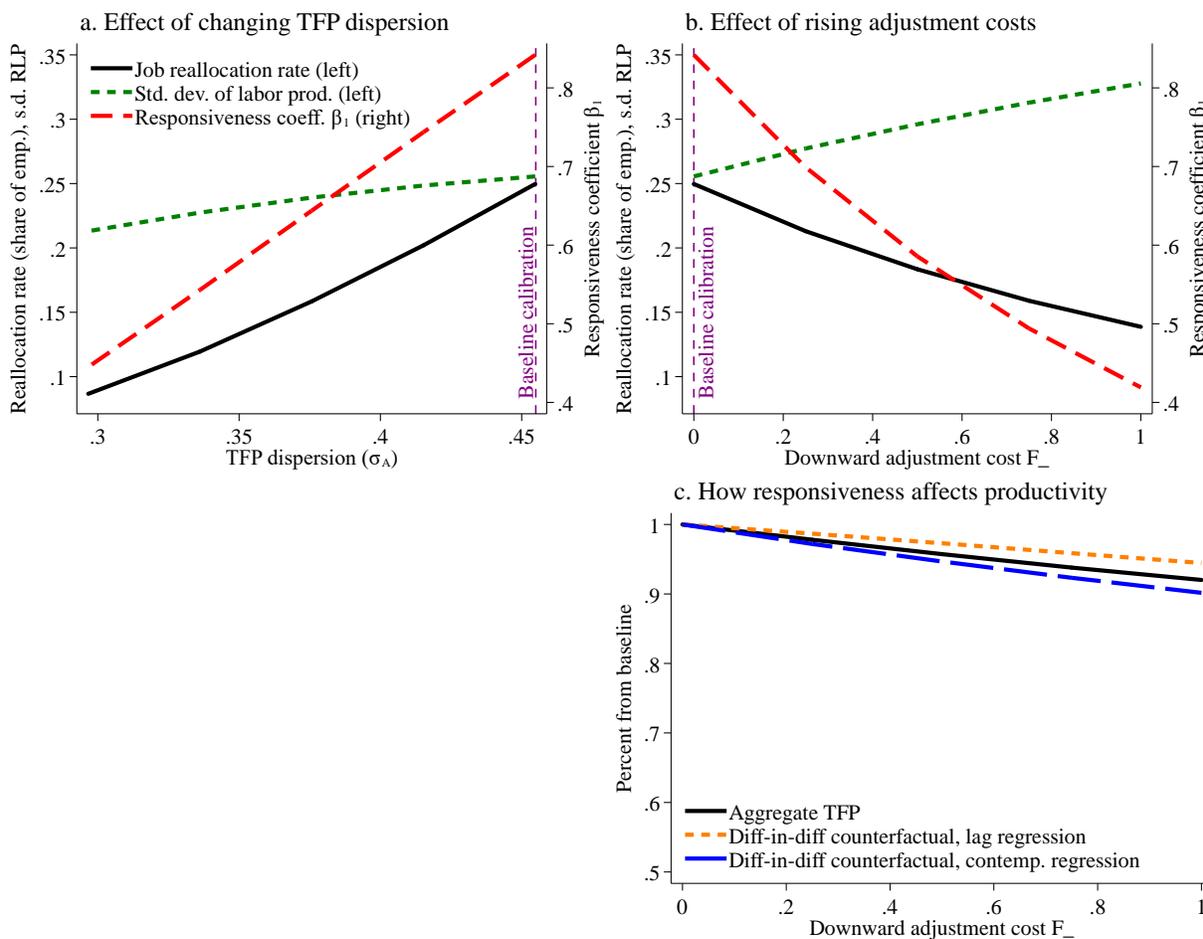
† Not statistically significant.

Figure 1: Job reallocation patterns differ by sector



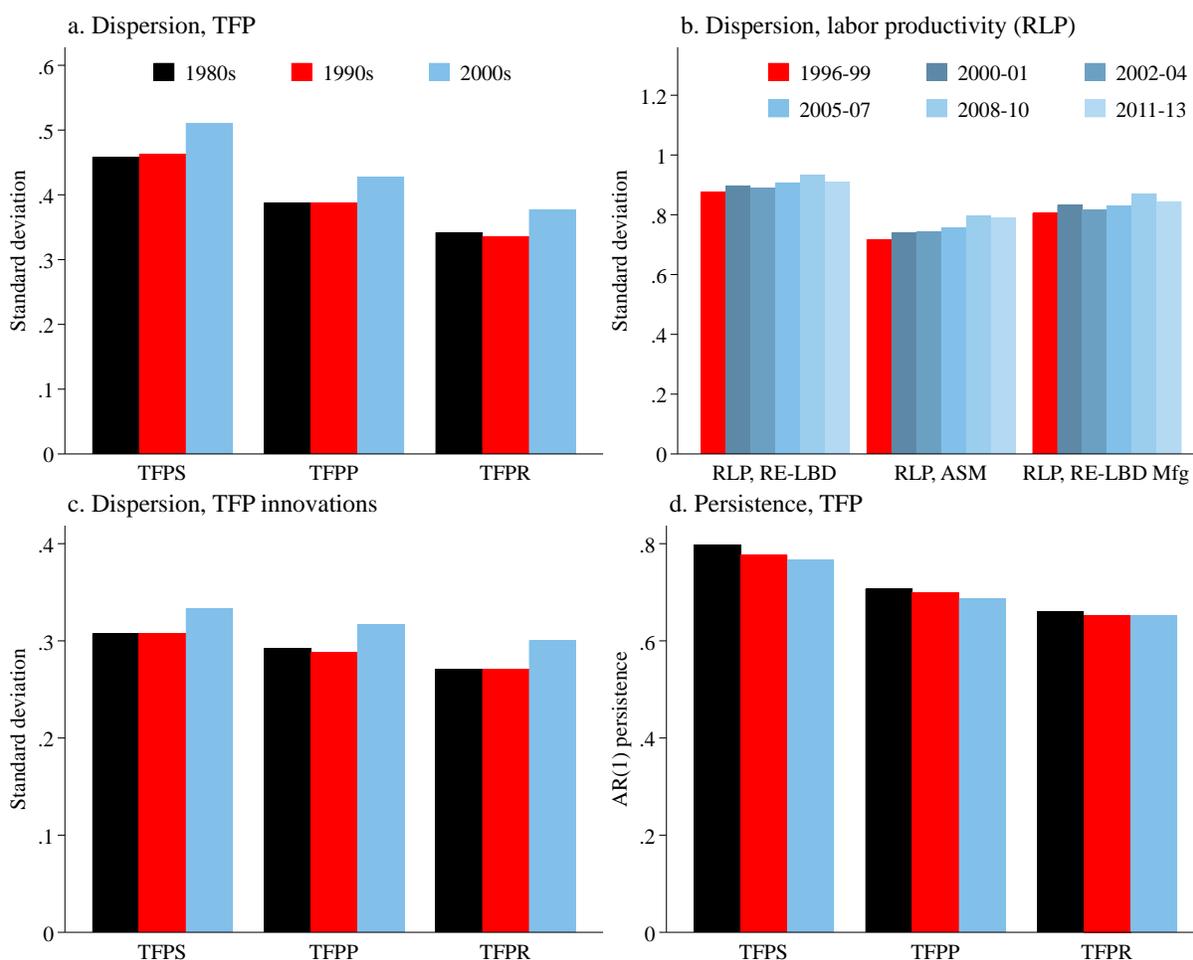
Note: HP trends using parameter set to 100. Industries defined on a consistent NAICS basis; high-tech is defined as in Hecker (2005). Data include all firms (new entrants, continuers, and exiters). Source: LBD.

Figure 2: The shocks and responsiveness hypotheses, model results



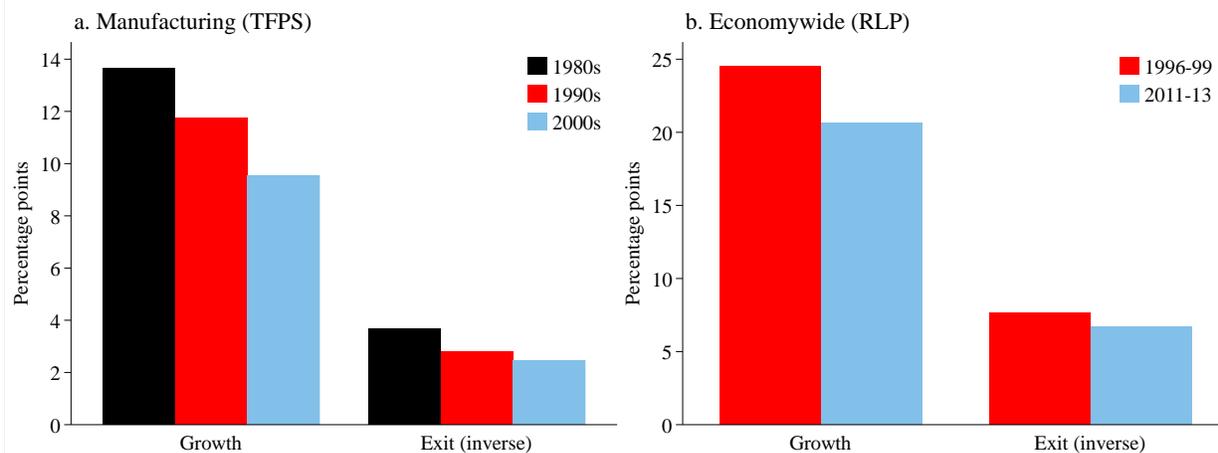
Note: Panels a and b share same legend. Results relative to model baseline calibration (vertical purple line) with downward adjustment cost $F_- = 0$ and TFP dispersion $\sigma_A = 0.46$ (see Appendix I and Table A1 for model calibration details). "s.d. RLP" refers to the standard deviation of revenue labor productivity in model-simulated data.

Figure 3: Within-industry productivity dispersion has risen



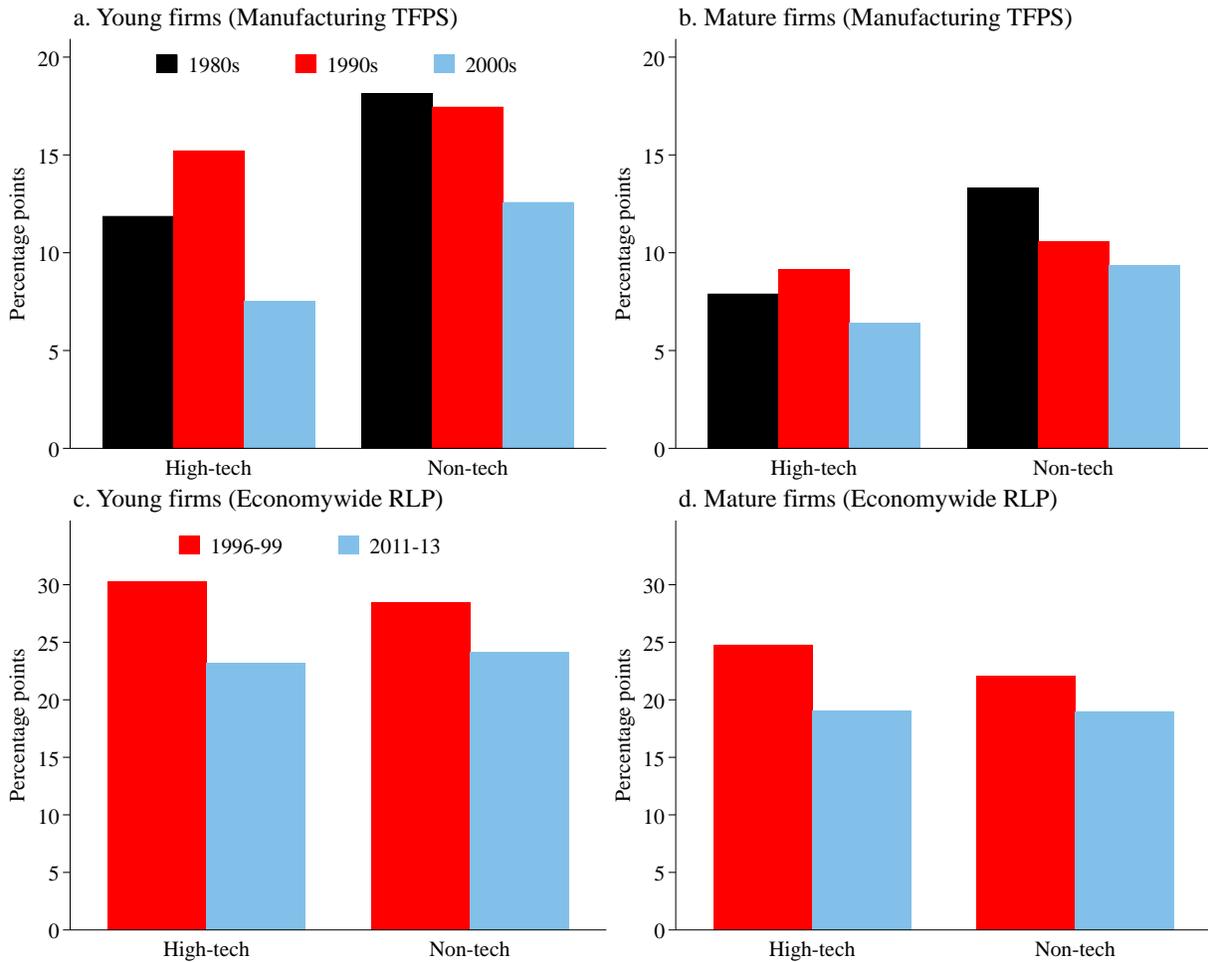
Note: Dispersion measures refer to standard deviation of within-industry (log) productivity. Panels a, c, and d share same legend. Persistence measures refer to AR(1) parameter. Source: ASM-CM (panels a, c, and d); RE-LBD (panel b).

Figure 4: Job growth and exit have become less responsive to productivity



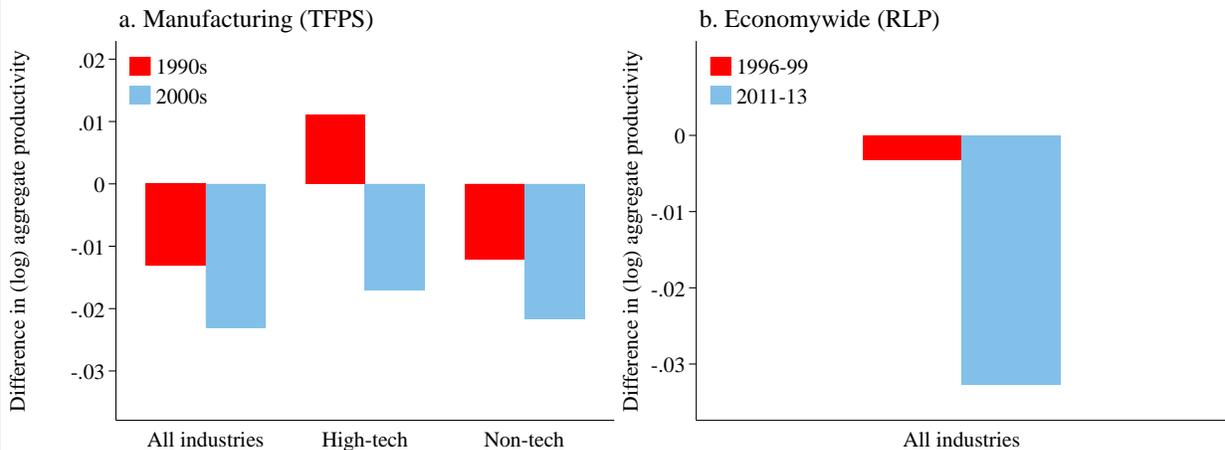
Note: Compares employment growth rate or (inverse) exit probability of establishment or firm that is one standard deviation above its industry-year mean productivity, versus the mean. Source: ASM-CM (panel a); RE-LBD (panel b).

Figure 5: Employment growth responsiveness: Young vs. mature firms, high-tech vs. non-tech



Note: Compares employment growth of establishment (panels a, b) or firm (panels c, d) that is one standard deviation above its industry-year mean productivity, versus the mean. Source: ASM-CM (panels a, b); RE-LBD (panels c, d).

Figure 6: Declining responsiveness and aggregate productivity



Note: Diff-in-diff counterfactual comparing model-predicted change in productivity from t to $t+1$ under constant responsiveness vs. actual responsiveness (see text). High-tech defined as in Hecker (2005). Source: ASM-CM (panel a); RE-LBD (panel b).

Appendix I: Model

A. Model environment.

Consider the following model of firm-level adjustment costs.⁵⁰ A firm maximizes the present discounted value of profits. The firm's value function and its components are specified as follows:

$$V(E_{jt-1}, A_{jt}) = \max \left\{ A_{jt} E_{jt}^\phi - W_t E_{jt} - C(H_{jt}, E_{jt-1}) + \beta \mathbb{E}_A(E_{jt}, A_{jt+1}) \right\} \quad (\text{A1})$$

with

$$C(H_{jt}, E_{jt-1}) = \begin{cases} \frac{\gamma}{2} \left(\frac{H_{jt}}{E_{jt-1}} \right)^2 + F_+ \max(H_{jt}, 0) + F_- \max(-H_{jt}, 0) & \text{if } H_{jt} \neq 0 \\ 0 & \text{otherwise} \end{cases}$$

where $\phi \leq 1$ due to product differentiation such that $A_{jt} E_{jt}^\phi$ is the revenue function for firm j , E_{jt} is employment for time t , H_{jt} is net hires made at the beginning of time t such that $H_{jt} = E_{jt} - E_{jt-1}$ (this can be positive or negative), W_t is the wage, and A_{jt} is a composite shock reflecting both technical efficiency and demand shocks. We interpret the revenue function curvature as reflecting product differentiation rather than decreasing returns to help draw out the relationship between revenue productivity and technical efficiency. That is, let firm-level price be given by $P_{jt} = D_{jt} Q_{jt}^{\phi-1}$, where $Q_{jt} = \tilde{A}_{jt} E_{jt}$ is firm-level output subject to a constant returns technology, with $A_{jt} = D_{jt} \tilde{A}_{jt}^\phi$. That is, A_{jt} is what we refer to as ‘‘TFP’’ in the main text, both in the conceptual framework and empirical analysis. Since labor is the only production factor, TFPR and revenue labor productivity (RLP) are both given by $P_{jt} \tilde{A}_{jt}$. Note that this specification nests the price-taking version of the model in which $\phi = 1$ as a special case, in which case TFP, TFPR, and RLP are equivalent. We focus on the $\phi < 1$ case in our calibration. We also abstract from demand shocks for clarity of exposition (i.e., $D_{jt} = 1 \forall j, t$).

This simple adjustment cost model is similar to Cooper, Haltiwanger, and Willis (2007, 2015), Elsby and Michaels (2013), and Bloom et al. (2018) and, in principle, accommodates both convex and non-convex adjustment costs. In particular, given the cost function $C(H_{jt})$, which

⁵⁰ We use the term ‘‘firm’’ for expositional purposes; in model exercises we do not distinguish between firms and establishments. Our empirical exercises using TFP measures and manufacturing data rely on establishments, while our economywide exercises using RLP rely on firms.

depends upon E_{jt-1} , the policy rule for hiring depends on the initial *state* faced by the firm, which is summarized as (E_{jt-1}, A_{jt}) .

We view the model as primarily illustrative but seek a reasonable baseline calibration that matches key features of the data and the parameters of the existing literature. Appropriate caution is needed since 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. We regard the calibration as providing guidance about the qualitative predictions for the key data moments we study but within a reasonable range of the parameter space.

Our main calibration exercise, described in detail below, implements “general equilibrium” in the sense that we fix the labor supply then find the wage that clears the labor market. Given a rigid labor supply, this may be thought of as an extreme scenario. However, in unreported exercises we 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 effect of changing responsiveness on aggregate productivity growth do not substantively depend on general versus partial equilibrium.

B. Model solution

Our method for solving the model is as follows. We create a state space for employment, with 2,400 points (distributed more densely at lower values), and for TFP realizations, with 115 points. We specify firm-level TFP to follow an AR(1) process, $\ln A_{jt} = \rho_a \ln A_{jt-1} + \eta_{jt}$, and in practice we use a Tauchen (1986) method for generating TFP draws. Table A1 reports our calibration choices, some of which are standard in the literature and others of which are designed to target specific data moments. We focus on kinked adjustment costs but report results below for convex adjustment costs as well. Empirically determined calibration choices are intended to produce a model economy that resembles the U.S. manufacturing sector in the 1980s, the initial timing of our empirical exercises, but the qualitative model results in which we are interested are not sensitive to these specific calibration choices.

We obtain policy functions using value function iteration then simulate 2,000 firms for 1,000 periods, jumping off from the stationary distribution of productivity but discarding the first

100 periods. Given a fixed (inelastic) labor supply, we check market clearing then adjust the wage using simple bisection until the labor market clears. We estimate responsiveness regressions and construct other statistics described in the text by using the simulated data generated by the model when in equilibrium.⁵¹

C. Main experiments

We perform several exercises on the model-simulated data with a focus on three key outcomes: aggregate job reallocation, the dispersion of revenue productivity (where in the model, revenue productivity is given by $A_{jt}E_{jt}^\phi$), and the responsiveness of growth to productivity as measured with the regression in equation (2) of the main text. That is, we measure the standard deviation of labor productivity in the model economy, and we estimate the following regression:

$$g_{jt+1} = \beta_0 + \beta_1 a_{jt} + \beta_2 e_{jt} + \varepsilon_{jt+1} \quad (\text{A2})$$

Where, as in the main text, g_{jt+1} is DHS employment growth from year t to year $t + 1$, a_{jt} is (initial) productivity, and e_{jt} is (initial) employment. This is the same as equation (2) and follows a timing convention that is analogous to our empirical work (though we confirm below that this timing convention is unimportant for the model’s qualitative results). “Responsiveness” is measured by β_1 .

We study labor productivity dispersion and responsiveness under two model experiments starting from the model’s baseline calibration. In our first experiment, we reduce the parameter governing true TFP dispersion, starting from its baseline calibrated value of $\sigma_a = 0.46$ (see Table A1). This is reported on Figure 2a in the main text. As TFP dispersion falls, aggregate job reallocation declines, labor productivity dispersion decreases, and responsiveness becomes weaker. This summarizes the “shocks” hypothesis: the declining pace of job reallocation we observe empirically could be explained by declining dispersion of TFP realizations if we were to also observe declining labor productivity dispersion and declining responsiveness. As shown in the main text, however, we actually observe rising labor productivity dispersion in our empirical work.

In our second experiment, we study the effects of rising adjustment costs. In particular, starting with the baseline calibration (where upward adjustment has a cost parameter of $F_+ = 1.51$) we raise the cost of downward adjustments (F_-). Figure 2b in the main text shows the

⁵¹ The code for solving our model can be found at http://rdecker.net/materials/DHJM2019_matlab.zip.

result of this experiment. Rising adjustment costs generate lower job reallocation, wider labor productivity dispersion, and weaker responsiveness, as we observe in our empirical exercises. This experiment suggests that rising adjustment costs—which we interpret broadly—can explain declining job reallocation. However, adjustment costs need not be the only explanation. Any changes to the costs or incentives to adjust employment can have these effects. The critical point is that we observe declining responsiveness alongside rising labor productivity dispersion.

The model results are robust to a wide range of conditions. Figure A1a shows that responsiveness regressions using concurrent TFP or RLP make the same qualitative predictions as regressions using lag TFP, as do regressions using the current or lagged TFP *innovation* (in the main text, we also find that our empirical results are robust to using innovations).

Figure A1b reports responsiveness coefficients from instrumental variables regressions performed on model-simulated data; these correspond with those we estimated on empirical data (described in the main text and Table A3). The solid red line reports our baseline OLS coefficient (matching Figure 2b in the main text), the short-dashed green line reports the responsiveness coefficient when the initial employment variable in the regression is instrumented with its own lag, and the long-dashed blue line reports the responsiveness coefficient when both initial employment and productivity are instrumented with their lags. Responsiveness declining with adjustment costs is robust to these specifications and closely tracks the OLS estimate. Figure A1c addresses the measurement error issue more specifically; we will describe that model exercise in more detail further below.

Figure A1d repeats our main responsiveness exercises in a model with convex (i.e., quadratic) adjustment costs rather than non-convex costs.⁵² This model’s predictions for responsiveness are the same as our model with non-convex costs. In short, the qualitative empirical predictions generated by our model are quite general.

D. Aggregate job reallocation implications

⁵² We construct an alternative baseline calibration of the model in which nonconvex costs are set to zero ($F_+ = F_- = 0$), but $\gamma = 3$ to again replicate a job reallocation rate of 0.25, leaving all other parameters unchanged relative to Table A1. (Recall from the model description that γ governs quadratic adjustment costs on employment). From this alternative convex cost baseline, we conduct both of our model experiments: (1) reduce TFP dispersion σ_a , and (2) raise adjustment cost γ above its baseline value. The qualitative results of the experiments for job reallocation, responsiveness, and revenue productivity dispersion are the same as those found in our non-convex cost experiments. The responsiveness results for the convex cost case are reported on Figure A1c.

We construct an estimate of the implied decline in job reallocation from the estimated decline in responsiveness as follows. We obtain a calibration of our model in which TFP dispersion is set to its 2000s value (rather than its 1980s value as in our baseline calibration; see Figure 3a) and the responsiveness coefficient β_1 is 30 percent weaker than the baseline calibration, consistent with the 30 percent decline in responsiveness from the 1980s to the 2000s found empirically (for TFPS; see Table 1 of the main text). This calibration produces a job reallocation rate of about 21 percent; in U.S manufacturing data, job reallocation averaged a bit under 22 percent during the 2000s (compared with 25 percent in the 1980s and in our baseline model calibration; see Figure 1). In other words, the model suggests that the estimated decline in responsiveness implies a decline in job reallocation broadly consistent with what we observe empirically over the same period.

E. Aggregate productivity implications

Define aggregate (sector-level) productivity as $A_t = Q_t/E_t$, that is, aggregate (sectoral) output divided by aggregate employment. Aggregate output using the CES aggregator underlying the firm-level demand structure specified above is given by:

$$Q_t = \left(\sum_j Q_{jt}^\phi \right)^{\frac{1}{\phi}}$$

where $\phi \leq 1$. Aggregate productivity follows:

$$A_t = \frac{Q_t}{E_t} = \frac{(\sum_j A_{jt} E_{jt}^\phi)^{\frac{1}{\phi}}}{E_t}$$

We show in Figure 3c that a rise in adjustment frictions from the baseline yields a decline in true aggregate productivity.⁵³ To provide an empirical framework for assessing effects on aggregate productivity, we also construct the following diff-in-diff object, described more fully in the main text:

$$\Delta_t^{t+1} = \sum_j \hat{\theta}_{jt+1}^L a_{jt} - \sum_j \hat{\theta}_{jt+1}^H a_{jt}$$

⁵³ That is, as in our main experiments, we start from the baseline calibration with $F_+ = 1.45$ and $F_- = 0$ then raise F_- .

where, as described in the main text, the $\hat{\theta}_{jt+1}^L$ and $\hat{\theta}_{jt+1}^H$ reflect the implied counterfactual employment distributions for each given responsiveness regime (low or high responsiveness, corresponding with $F_- > 0$ and $F_- = 0$ in our model, respectively). On Figure 2c in the main text, we report this diff-in-diff object for several “high-cost” scenarios relative to the “low-cost” baseline. We construct the object under two different timing assumptions (see main text), and the resulting diff-in-diff objects closely track true productivity effects of declining responsiveness.

This diff-in-diff counterfactual isolates the impact of changing responsiveness on the weighted productivity mean holding everything else constant. In particular, it implicitly relies on estimated policy functions that reflect optimal growth responses of firms to their productivity draws and initial size, taking the degree of revenue curvature as given.

F. Alternative framework: Wedges

In the discussion above we adopted an adjustment cost model in which is it straightforward to consider theoretical explanations for empirical changes in productivity dispersion and responsiveness. However, while we find the adjustment cost framework convenient, we do not argue that it is the only (or even preferred) approach. As an example, here we show how a broader interpretation can be adopted, following the seminal work of Hsieh and Klenow (2009).

Hsieh and Klenow (2009) show how measured revenue productivity dispersion can exist in equilibrium if there are “distortions” or wedges affecting firms’ first-order conditions. This framework can be viewed as a reduced form way of capturing many factors. In principle, these wedges could be broadly interpreted as a proxy for the adjustment frictions discussed above. However, the typical approach is to model these wedges in a simple static framework. Specifically, consider a simple one-factor (employment) model in the spirit of Hsieh and Klenow (2009). Firms maximize period t profits given by:

$$S_{jt}A_{jt}E_{jt}^\phi - W_tE_{jt} \tag{A3}$$

where $A_{jt}E_{jt}^\phi$ is revenue and S_{jt} is a firm-specific wedge, which can be thought of as a tax when $S_{jt} < 1$ or as a subsidy when $S_{jt} > 1$. Then the first-order condition is given by:

$$E_{et} = \left(\frac{\phi S_{jt} A_{jt}}{W_t} \right)^{\frac{1}{1-\phi}}. \tag{A4}$$

Taking logs (indicated by lower case) and time differences (indicated by Δ) and sweeping out year and industry effects yields a kind of firm-level growth rate:

$$\Delta e_{jt} = \frac{1}{1-\phi} (\Delta s_{jt} + \Delta a_{jt}). \quad (\text{A5})$$

Intuitively, employment growth is determined by the change in the wedge and the change in the productivity realization (as well as revenue function curvature, where more curvature from lower ϕ dampens growth).

Suppose that wedges are correlated with fundamentals such that they follow this process:

$$s_{jt} = -\vartheta a_{jt} + v_{jt}, \quad (\text{A6})$$

where, consistent with much of the recent literature, we assume $\vartheta \in (0,1)$, and v_{jt} is independent of a_{jt} with $\mathbb{E}(v_{jt}) = 0$.⁵⁴ Given logs, $s_{jt} < 0$ is consistent with $S_{jt} < 1$. Equation (A6) states that firms with more favorable fundamentals (e.g., higher TFP) face greater wedges on average (by this we mean lower S_{jt}), but the variance of (log) distortions is lower than the variance of fundamentals. Combining (A6) and (A5) yields:

$$\Delta e_{jt} = \frac{1}{1-\phi} ((1-\vartheta)\Delta a_{jt} + \Delta v_{jt}) \quad (\text{A7})$$

Equation (A7) shows that the relationship between employment growth and productivity depends not only on ϕ but also on ϑ , which determines the covariance between firm productivity and firm distortions. A higher value of ϑ results in weaker responsiveness of growth to productivity. In the text, we refer to a higher ϑ as reflecting a more positive correlation between fundamentals and distortions. By this we mean that the implicit tax on firms is increasing in fundamentals. In this case the implicit tax is larger the less positive is s_{jt} . Note also that aggregate job reallocation, which in this context can be thought of as the dispersion of employment growth rates, is decreasing in ϑ .

This framework also has implications for revenue productivity dispersion. Log revenue per worker is given by $w - \ln \phi + \vartheta a_{jt} - v_{jt}$, such that the dispersion of revenue labor productivity is *increasing* in ϑ .

⁵⁴ By “consistent with the literature,” we mean a common finding in the literature is that indirect measures of wedges (i.e., revenue productivity measures like TFPR) are positively correlated with measures of fundamentals (technical efficiency and demand shocks) and have lower variance than fundamentals. See Foster, Haltiwanger, and Syverson (2008) and Blackwood et al. (2019).

This model, albeit highly simplified, thus yields rich empirical predictions. In particular, suppose there is an increase in ϑ . Then:

1. Responsiveness—the relationship between employment growth and firm productivity shocks—will weaken
2. Job reallocation will decline
3. Revenue productivity dispersion will rise

That is, this model can generate a decline in reallocation and responsiveness along with a rise in revenue productivity dispersion if distortions become larger as productivity strengthens; all the predictions for rising adjustment frictions also hold for this reduced form model of wedges.

There are some differences, however, between the adjustment cost model and the model with static distortions. The adjustment cost model is explicitly dynamic and results in a policy function that relates growth to the *level* of a_{jt} while controlling for the state, given by e_{jt-1} ; the model of static distortions instead generates a policy rule that depends on Δa_{jt} and is independent of e_{jt-1} . That said, we do perform empirical exercises that are consistent with the static model, where we relate employment growth to productivity innovations instead of levels. These are given in Table 2 of the main text.

The wedge model also yields similar but slightly different implications for changes in the variance of shocks. A decline in the variance of a_{jt} yields declining reallocation and revenue productivity dispersion but does not yield a change in marginal responsiveness.

Relatedly, within the main adjustment cost framework, a change in the curvature of the revenue function (determined by the parameter ϕ , where the inverse elasticity of demand is $\phi - 1$) can also affect responsiveness and reallocation and, in the presence of plausible adjustment costs, the dispersion of labor productivity.⁵⁵ Figure A2 reports the results of starting from the baseline model calibration and reducing the parameter ϕ ; panel (a) employs the baseline model with non-convex (“kinked”) adjustment cost (i.e., the model used for the main simulation results in the paper, as on Figure 2), and panel (b) employs the model calibrated for non-convex costs (i.e., $\gamma = 3$).⁵⁶ In both panels, the far right side is the baseline calibration (in which job

⁵⁵ Note, however, that in the version of the model without wedges or frictions, labor productivity dispersion is zero regardless of the value of ϕ .

⁵⁶ As noted previously we construct an alternative baseline calibration of the model in which non-convex costs are set to zero ($F_+ = F_- = 0$), but $\gamma = 3$ to again replicate a job reallocation rate of 0.25, leaving all other parameters unchanged relative to Table A1.

reallocation is set to 25 percent). Moving left from the baseline, we reduce the value of ϕ , which adds curvature to the revenue function, while leaving costs constant at their baseline value. This can be thought of as exogenously increasing the markup (going from the right side to the left side of the figure panels). In both of these specific settings (a non-convex cost economy and a convex cost economy), an increase in revenue function curvature coming from a reduction in ϕ generates lower responsiveness and reallocation. However, the effects on labor productivity dispersion are modest and model dependent: increased revenue function curvature raises labor productivity dispersion in the presence of non-convex costs but reduces labor productivity dispersion (modestly) in the presence of convex costs. In broad terms this can also be thought of as a special case of the general reduced form wedge model and highlights that a number of forces can cause lower responsiveness, though the implications for revenue productivity dispersion are theoretically ambiguous.

An important feature not incorporated in the exercises of Figure A2 is a rise in the dispersion of idiosyncratic variable markups, potentially with an increase in the correlation of markups with fundamentals. Such a change would be consistent with an increase in idiosyncratic wedges that are correlated with fundamentals.

G. Measurement error

Returning to the baseline adjustment cost model, we can study the consequences of error in the measurement of employment. We consider three possible scenarios. First, the level of employment (as related to the coefficient β_2 and the dependent variable g_{jt+1} in equation (A2) above) is measured with error, but the employment variable used to calculate productivity is not. Second, the employment variable used to calculate productivity is measured with error, but employment measures elsewhere are not.⁵⁷ Third, both employment and productivity are measured with error.

We run our adjustment cost experiment—that is, we increase the cost of upward employment adjustment from the baseline—under the three measurement error scenarios above. We add error to econometrician-observed employment in each year by multiplying true firm-level employment by a disturbance term drawn from an independent normal distribution with

⁵⁷ Recall that in our manufacturing regressions, the employment variable used to measure productivity, which comes from the ASM/CM, is independent of the employment variable used to measure employment levels and growth, which comes from the LBD.

mean 1 and standard deviation 0.033, such that employment disturbances of 10 percent map to three standard deviations from the mean.⁵⁸ In each case, the firm observes its true employment, but the econometrician observes employment with error. We then estimate equation (A2) under each measurement error scenario.

Figure A1c reports the result of these measurement experiments along with our error-free experiments. No form of measurement error changes the qualitative pattern of responsiveness that declines as adjustment costs rise. Measurement error in *either* productivity *or* employment levels (and growth rates) shifts all coefficients down across the range of adjustment costs. Measurement error in *both* productivity *and* employment shifts all coefficients up across the range of adjustment costs. Notably, however, the differences between the error-free coefficients and the error-based coefficients are very small. Differences in coefficients large enough to explain our results would require employment measurement error much larger than the error we specify (in which disturbances of 10 percent map to three standard deviations from the truth).

Given constant measurement error, then, our empirical results will overstate the relationship between productivity and growth if both employment and productivity are measured with error, or our results will understate the relationship between productivity and growth if either employment or productivity (but not both) is measured with error. The downward trend in responsiveness we document empirically could be explained by measurement error if certain conditions hold. For example, if only employment or productivity (but not both) is measured with error, rising measurement error over time could appear empirically as declining responsiveness. Alternatively, if both employment and productivity are measured with error, then diminishing measurement error over time could appear empirically as declining responsiveness. In either case, the level and change in error would have to be very large to be explanatory. We are not aware of any evidence that our data sources have seen these particular patterns of measurement error over time.

⁵⁸ This choice is arbitrary and does not have qualitative implications.

Appendix II: Data

A. Longitudinal Business Database

For longitudinal information we rely on the Longitudinal Business Database (LBD), which covers the universe of private nonfarm employer business establishments in the U.S. The LBD records establishment employment, payroll, detailed industry, and location annually (with employment corresponding to March 12). Establishments are linked over time by high-quality longitudinal identifiers, and firm identifiers link establishments of multi-establishment firms. See Jarmin and Miranda (2002) for a description of the LBD, which is constructed from the Census Bureau's Business Register. The LBD's high-quality longitudinal linkages make it ideal for studying growth and survival outcomes of businesses.

In our regression specifications we include several establishment characteristic controls derived from the LBD. Key among them is firm age. We follow the large LBD-based literature in defining firm age as follows. Upon the first appearance of a firm identifier in the LBD, we define firm age as the age of the firm's oldest establishment, where an establishment has age 0 during the year in which it first reports positive payroll. Thereafter, the firm ages naturally (i.e., we add one year to the firm's age for each calendar year after the firm identifier's first observation). This allows us to abstract from spurious changes in firm identifiers. We also use firm identifiers to measure firm size, which is the sum of employment across all the firm's establishments. In our regressions we control for firm size based on four cutoffs: fewer than 250 employees, 250-499 employees, 500-999 employees, and 1,000 or more employees (these cutoffs follow Foster, Grim, and Haltiwanger (2016), hereafter FGH).

B. Revenue-enhanced LBD (RE-LBD)

While the LBD does not include revenue data, revenue is available in the Business Register at the employer identification number (EIN) level starting in the mid-1990s. Importantly, EINs are not a straightforward firm or establishment identifier in that multiple establishments can have the same EIN, and some firms can have multiple EINs (e.g., splitting the firm by geography or separating tax functions from payroll functions). In the case of multi-establishment firms, in general revenue data are not broken out by establishment. Haltiwanger et al. (2017) deal with these various challenges and create firm-level revenue data by aggregating

across EINs of the same firm.⁵⁹ They then match these revenue figures to the LBD at the firm level, finding nominal revenue figures for about 80 percent of LBD firms. The resulting revenue dataset is roughly representative of the overall LBD in terms of observables like firm age, firm size, sector, multi- or single-establishment status, and patterns of firm growth. Nevertheless, Haltiwanger et al. (2017) construct propensity scores for the entire LBD using logistic regressions with dependent variable equal to 1 for firms with revenue data and 0 otherwise. These regressions are run separately for birth, deaths, and continuers, and they rely on observables including firm size, firm age, employment growth rate, industry, and multi-establishment status. We use the resulting propensity scores (in inverse) as sampling weights in all regressions. We deflate revenue with the GDP deflator, but this is unimportant as all empirical exercises will implicitly control for industry-level prices as we deviate firm productivity from industry-year means.

C. Manufacturing data

We supplement the LBD with manufacturing data from the Census of Manufacturers (CM) and the Annual Survey of Manufacturers (ASM), a dataset we obtain from FGH and update through 2013. The CM surveys almost the universe of manufacturing establishments every five years (those ending in “2” and “7”); we use CM data from 1982 through 2012.⁶⁰ The ASM, conducted in non-CM years, surveys roughly 50,000-70,000 establishments; we use ASM data from 1981 through 2013. The ASM is a series of five-year panels (starting in years ending in “4” and “9”) with probability of panel selection being a function of industry and size.

We combine the CM and ASM into an annual manufacturing establishment dataset covering 1981-2013, and we link the combined ASM-CM with the LBD by establishment and year using internal Census Bureau establishment identifiers that are consistent across these datasets. We then create a dummy variable equal to 1 for those establishments that appear in both the ASM-CM and the LBD and 0 for those establishments that appear only in the LBD. We then create propensity scores using a logistic regression to predict ASM-CM presence based on the following variables: whether the establishment is part of a multi-establishment firm, size

⁵⁹ This is a complicated process involving careful attention to details including industry and legal form of organization, which can affect the way in which revenue data are reported and the way EINs map to firms.

⁶⁰ Very small establishments (those with fewer than five employees) are not surveyed by the CM; the Census Bureau fills in data for these with administrative records. We do not include these cases.

(employment), payroll, detailed industry, and firm age. We estimate these propensity scores separately for each year; we then use them (in inverse) as sampling weights in all regressions.

D. Output and production factors

We require measures of revenue and production factors to construct TFPS, TFPP, and TFPR. We calculate real establishment-level revenue (or, under TFPR assumptions, output) as $Q_{jt} = (TVS_{jt} + DF_{jt} + DW_{jt})/PISHIP_t$, where TVS_{jt} is total value of shipments, DF_{jt} is the change in (the value of) finished goods inventories, DW_{jt} is the change in (the value of) work-in-progress inventories, and $PISHIP_t$ is the *industry-level* shipments deflator, which varies by detailed industry (4-digit SIC prior to 1997 and 6-digit NAICS thereafter) and is taken from the NBER-CES Manufacturing Productivity Database. If the resulting Q_{jt} is not greater than zero, then we simply set $Q_{jt} = TVS_{jt}/PISHIP_t$.

For the purposes of TFP estimation, we construct labor from the ASM in terms of total hours (TH_{jt}) as follows:

$$TH_{jt} = \begin{cases} PH_{jt} \frac{SW_{jt}}{WW_{jt}} & \text{if } SW_{jt} > 0 \text{ and } WW_{jt} > 0 \\ PH_{jt} & \text{otherwise} \end{cases} \quad (\text{A8})$$

where PH_{jt} is production worker hours, SW_{jt} is total payroll, and WW_{jt} is the payroll of production workers.

We measure capital separately for structures and equipment using the perpetual inventory method: $K_{jt+1} = (1 - \delta_{t+1})K_{jt} + I_{jt+1}$ where K is the capital stock, δ is a year- (and industry-) specific depreciation rate, and I is investment. At the earliest year possible for a given establishment, we initialize the capital stock by multiplying the establishment's reported book value by a ratio of real capital to book value of capital derived from BEA data (where the ratio varies by 2-digit SIC or 3-digit NAICS). Thereafter, we observe annual capital expenditures and update the capital stock accordingly, where we deflate capital expenditures using BLS deflators.⁶¹

We calculate materials as $M_{jt} = (CP_{jt} + CR_{jt} + CW_{jt})/PIMAT_t$, where CP is the cost of materials and parts, CR is the cost of resales, CW is the cost of work done for the establishment (by others) on the establishment's materials, and $PIMAT$ is the industry materials deflator. We

⁶¹ See FGH for more detail. In a small number of cases (less than 0.5 percent) we cannot initialize the capital stock as described; in such cases we follow Bloom et al. (2013) using I/K ratios.

calculate energy costs as $E_{jt} = (EE_{jt} + CF_{jt})/PIEN_t$, where EE is the cost of purchased electricity, CF is the cost of purchased fuels consumed for heat, power, or electricity generation, and $PIEN$ is the industry energy deflator.

We use the production factor and output measures described above for each of our three TFP measures (TFPS, TFPP, and TFPR).

E. Cost and revenue shares: TFPS and TFPR

TFPS and TFPR productivity estimates require industry-level factor expenditures as shares of revenue (for TFPS) or cost (for TFPR) to construct factor elasticity estimates. We obtain these shares at the detailed industry level (4-digit SIC prior to 1997, 6-digit NAICS thereafter) from the NBER-CES Manufacturing Productivity Database, which reports industry-level figures for expenditures on equipment, structures, materials, energy, and labor. We average these cost shares across all of 1981-2013 to obtain time-invariant elasticities, though our results are robust to instead using time-varying elasticities as in FGH.

F. Proxy method: TFPP

Our TFPP productivity concept requires us to estimate factor elasticities using proxy methods. Given the challenge of identifying exogenous shocks to fundamentals, a long literature (Olley and Pakes (1996), Levinsohn and Petrin (2003)) proposes using a variable production factor as a “proxy” for identification. Blackwood et al. (2019) compare multiple proxy-based TFP concepts with other concepts from the literature. Some literature achieves this using a two-step procedure (see Akerberg, Caves, and Frazer (2015)), but we follow Wooldridge (2009) in implementing a single-step GMM approach using lagged values of capital and variable inputs as instruments. We refer the reader to the just-mentioned research for more detail on the general approach to proxy estimation of production functions. For our purposes, we estimate factor elasticities separately by 2- and 3-digit industries using energy as the proxy variable.

Appendix III: Additional empirical results

A. Reallocation has declined within firm age bins

As noted in the text, the aggregate decline in job reallocation is not simply a composition effect due to declining young firm activity. Rather, we also observe declining reallocation within firm age bins. To see this, we first create seven firm age groups (ages 0, 1, 2, 3, 4, 5 and 6+). We then study the change in aggregate (weighted average) job reallocation in year t relative to a base year t_0 with the following shift-share decomposition:

$$R_t - R_{t_0} = \sum_a \omega_{at_0}(R_{at} - R_{at_0}) + \sum_a R_{at_0}(\omega_{at} - \omega_{at_0}) + \sum_a (R_{at} - R_{at_0})(\omega_{at} - \omega_{at_0})$$

where R_t is the aggregate (or, as we will specify it below, sector-level) job reallocation rate, a indexes age bins, ω_{at} is the employment share of age group a in time t , and R_{at} is the reallocation rate for age group a in time t . The first term is a within-age-group component based on the change in flows among firms of that age. The second term is a between-group component capturing the change in the age composition. The third term is a cross term. We focus on the overall component and the within component; the residual coming from composition shifts and cross terms reflects the extent to which composition effects account for the aggregate change.

To abstract from business cycle issues, we construct this counterfactual between the business cycle peaks of 1987-1989, 1997-1999, and 2011-2013. We study the long differences in reallocation rates between these three periods. Figure A3 illustrates the results, showing both the overall change in reallocation for a sector and the change in the within-age-group term, indicated by the “Holding age constant” bars. As is evident, the decline in reallocation within age groups explains the bulk of the overall decline. In other words, the changing age composition of U.S. firms resulting from changing patterns of firm entry does not explain the patterns of reallocation that motivate this paper.

B. Further results on responsiveness to productivity innovations

The main text describes baseline regressions relating employment growth to the *innovation* to productivity for firms of all ages (reported on Table 2). Table A2 reports these same regressions (using innovations to TFPS) separating young and mature firms; the results are broadly consistent with our other responsiveness estimates.

C. Instrumental variables: empirical results

A particular challenge for our empirical approach is that our workhorse regression given by equation (12) in the main text features initial employment (E_{jt-1}) on the right-hand-side (as the state variable) and on the left-hand-side (in the DHS growth dependent variable). Additionally, in our economywide regressions using labor productivity, initial employment also appears in the denominator of the productivity indicator (which is real revenue per worker).⁶² In Appendix I, we explore this problem by running instrumental variables regressions on model-simulated data. Regressions in which an employment lag is used to instrument for initial employment (i.e., use E_{jt-2} as an instrument for E_{jt-1}), and regressions in which we additionally instrument for productivity using a lag, find that responsiveness still declines as adjustment costs rise. This suggests that we can evaluate robustness of our main responsiveness results to the employment endogeneity issue using similar instrumental variables regressions.

For brevity, we focus on the time-trend regression specifications for studying changing responsiveness. Table A3 reports results of instrumental variables regressions. The first column reports establishment-based results for the manufacturing sector using our preferred productivity measure, TFPS, and instrumenting for initial employment. The second column reports economywide firm-based results instrumenting for initial employment, and the third column reports economywide firm-based results instrumenting for initial employment and for productivity. In each column, and for both young and mature firms, we observe declining responsiveness as indicated by the negative (and statistically significant) coefficient on the linear trend variables.

⁶² Initial employment is also used in TFP estimation in our manufacturing-only exercises; however, the employment variable used for TFP is independently constructed from our ASM-CM dataset (see Appendix II).

Appendix references

- Akerberg, Daniel A., Kevin Caves, and Garth Frazer. 2015. "Identification Properties of Recent Production Function Estimators." *Econometrica* 83 no. 6:2411-2451.
- Blackwood, Glenn, Lucia Foster, Cheryl Grim, John Haltiwanger and Zoltan Wolf. 2019. "Macro and Micro Dynamics of Productivity: From Devilish Details to Insights", Working Paper.
- Bloom, Nicholas, Erik Brynjolfsson, Lucia Foster, Ron Jarmin, Itay Saporta-Eksten, and John Van Reenen. 2013. "Management in America." CES-WP-13-01, CES Discussion Paper Series, Washington, DC.
- Bloom, Nicholas, Max Floetotto, Nir Jaimovich, Itay Saporta-Eksten, and Stephen J. Terry. 2018. "Really Uncertain Business Cycles." *Econometrica* 86 no. 3:1031-65.
- 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 no. 1:37-50.
- Elsby, Michael W. L. and Ryan Michaels. 2013. "Marginal Jobs, Heterogeneous Firms, and Unemployment Flows." *American Economic Journal: Macroeconomics* 5 no. 1:1-48.
- Foster, Lucia, Cheryl Grim, and John Haltiwanger. 2016. "Reallocation in the Great Recession: Cleansing or Not?" *Journal of Labor Economics* 34 no. S1 (Part 2, January 2016):S293-S331.
- Hsieh, Chang-Tai and Peter J. Klenow. 2009. "Misallocation and Manufacturing TFP in China and India." *Quarterly Journal of Economics*, 124(4): 1403-48.
- 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.
- Haltiwanger, John, Ron S. Jarmin, Robert Kulick and Javier Miranda. 2017. "High Growth Young Firms: Contribution to Job Growth, Revenue Growth and Productivity" In *Measuring Entrepreneurial Businesses: Current Knowledge and Challenges* edited by John Haltiwanger, Erik Hurst, Javier Miranda, and Antoinette Schoar. Chicago: 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.
- Levinsohn, James and Amil Petrin. 2003. "Estimating Production Functions Using Inputs to Control for Unobservables." *Review of Economic Studies* 70 no. 2:317-41.
- Olley, G. Steven and Ariel Pakes. 1996. "The Dynamics of Productivity in the Telecommunications Industry." *Econometrica*, 64 no. 6:1263-98.
- Tauchen, George. 1986. "Finite State Markov-Chain Approximations to Univariate and Vector Autoregressions." *Economics Letters* 20 no. 2:177-181.
- Wooldridge, Jeffrey M. 2009. "On Estimating Firm-Level Productivity Functions Using Proxy Variables to Control for Unobservables." *Economics Letters* 104 no. 3:112-4.

Table A1: Baseline model calibration

	Description	Value	Calibration rationale
ϕ	Inverse demand elasticity parameter	0.80	Standard in literature
β	Discount factor	0.96	Standard in literature
ρ_a	(Log) TFP AR(1) coefficient	0.80	Estimated TFPS AR(1), 1980s average
σ_a	Standard deviation of (log) TFP	0.46	Estimated TFPS standard deviation, 1980s average
σ_η	Standard deviation of TFP innovation	0.28	Implied by ρ and σ_a
F_+	Upward kinked adjustment cost	1.51	Target job reallocation rate = 0.25 (1980s average)*
F_-	Downward kinked adjustment cost	0.00	(Rely on upward cost for baseline calibration)
γ	Quadratic adjustment cost	0.00	(Rely on non-convex costs for baseline calibration. In alternative baseline described in appendix, set $\gamma = 3$ and $F_+ = F_- = 0$.)

Moment targets refer to U.S. manufacturing sector.

*1980s average reallocation rate as shown on Figure 1 in the main text.

Table A2: Instrumental variables regressions, employment growth responsiveness

	TFPS: IV for employment	RLP: IV for employment	RLP: IV for emp & RLP
Prod*Young: β_1^y	0.4358 (0.0177)	0.3170 (0.0013)	0.1499 (0.0016)
Prod*Young*trend: δ^y	-0.0042 (0.0009)	-0.0033 (0.0001)	-0.0015 (0.0001)
Prod*Mature: β_1^m	0.3123 (0.0104)	0.2581 (0.0010)	0.1092 (0.0001)
Prod*Mature*trend: δ^m	-0.0020 (0.0005)	-0.0032 (0.0001)	-0.0014 (0.0001)
Observations (thousands)	2,179	4,909	4,909

Note: All coefficients statistically significant with $p < 0.01$. All regressions include controls described in equation (12) and related text. RLP regressions use 10 percent random sample of RE-LBD.

Figure A1: Alternative responsiveness coefficient β_1 specifications in model-simulated data

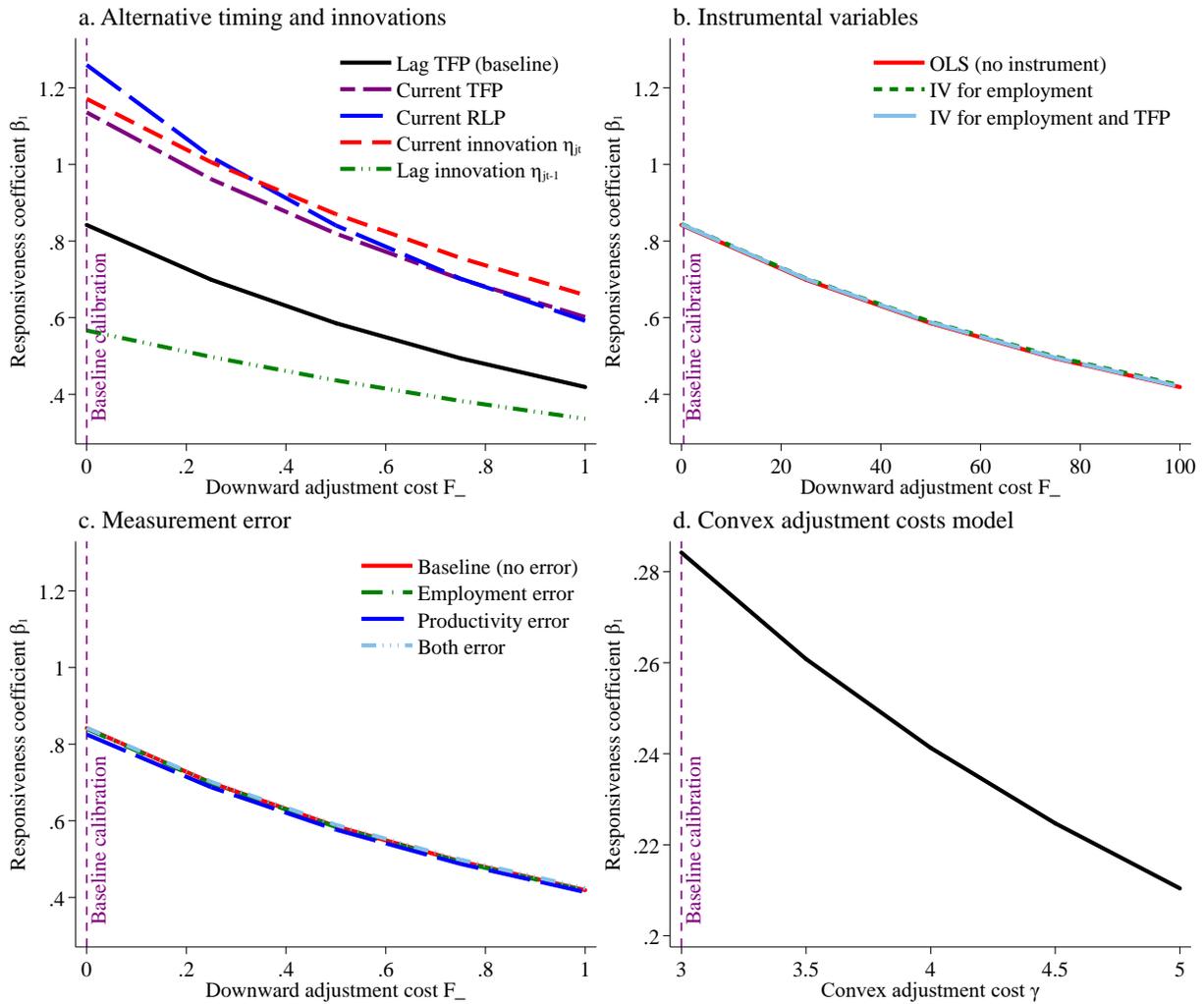
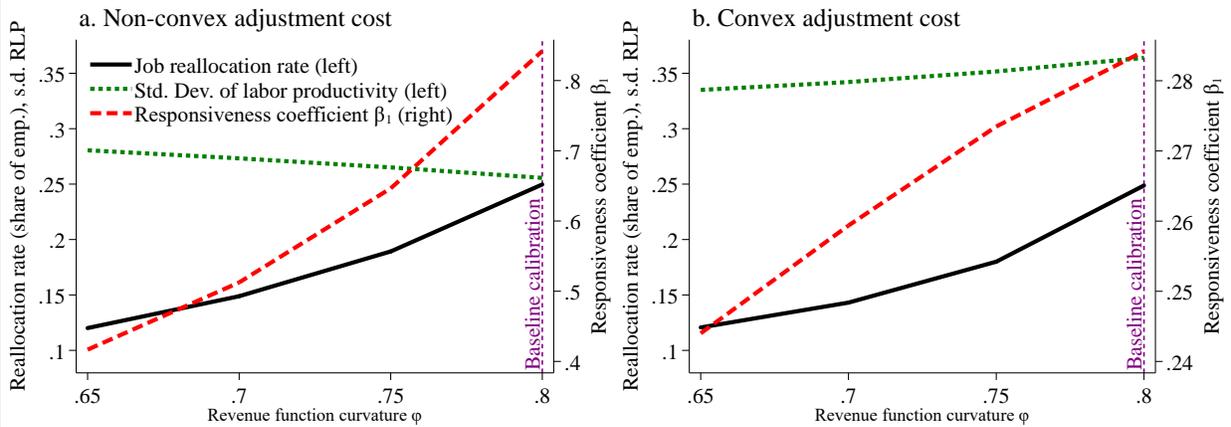


Figure A2: Effects of changing revenue function curvature



Note: Left panel refers to baseline model with kinked (upward) adjustment cost $F_+=1.45$. Panel (b) refers to model with convex (quadratic) adjustment cost $\gamma=3$. See model description in Appendix I for explanation of revenue function curvature parameter $\varphi-1$ denotes the inverse elasticity of demand.

Figure A3: Most variation in job reallocation is within firm age classes

