

Declining Responsiveness at the Establishment Level: Sources and Productivity Implications

Russell Cooper, John Haltiwanger, and Jonathan Willis

Working Paper 2024-3
February 2024

Abstract: This paper studies competing sources of declining dynamism. Evidence shows that an important component of this decline is accounted for by the reduction in the response of employment to shocks in US establishments. Using a plant-level dynamic optimization problem as a framework for analysis, four potential reasons for this decline are studied: (i) a change in exogenous processes for profits, (ii) an increase in impatience, (iii) increased market power, and (iv) increasing adjustment costs. We identify and quantify the contribution of each of these factors building on a simulated method of moments estimation of our structural model. Our results indicate that the reduction in responsiveness largely reflects increased costs of employment adjustment. Changes in market power, as captured by changes in the curvature of the revenue function, play a minimal role. But, in the presence of rising adjustment costs, measured sales-weighted markups using the recently popular *indirect production* approach rise substantially, along with rising dispersion and skewness of such measured markups.

JEL classification: E24, E32, J23

Key words: declining dynamism, adjustment costs, employment

<https://doi.org/10.29338/wp2024-03>

Excellent research assistance from Tuna Dökmeci and Cody Tuttle and financial support from the Kauffman Foundation are greatly appreciated. The authors thank Shawn Klimek, Kirk White, participants at the Society of Economic Dynamics Conference in 2017, the Midwestern Macro Conference in 2018, the ASSA Meetings in 2020, and the Technology and Declining Economic Dynamism in 2020 for helpful comments. The views expressed herein are solely those of the authors and do not necessarily reflect the views of the Federal Reserve Banks of Atlanta or Kansas City or the US Bureau of the Census. All results have been reviewed to ensure that no confidential information is disclosed. John Haltiwanger was a Schedule A employee at the US Bureau of the Census when this paper was written.

Please address questions regarding content to Russell Cooper, Department of Economics, the European University Institute, russellcoop@gmail.com; John Haltiwanger, Department of Economics, University of Maryland, haltiwang@econ.umd.edu; or Jonathan L. Willis, Research Department, Federal Reserve Bank of Atlanta, jon.willis@atl.frb.org.

Federal Reserve Bank of Atlanta working papers, including revised versions, are available on the Atlanta Fed's website at www.frbatlanta.org. Click "Publications" and then "Working Papers." To receive e-mail notifications about new papers, use frbatlanta.org/forms/subscribe.

1 Motivation

The decline in dynamism in US establishments is well documented. In the 1980s, the pace of job reallocation across establishments in the US private non-farm sector averaged 33.3 percent on an annual basis; in the post 2000 period, job reallocation averaged only 27.3 percent.¹ This reflects a decline of 20 percent in the pace of reallocation. Another indicator of the decline in dynamism is that the responsiveness of incumbent establishments to profitability shocks has declined over this period. This decline in responsiveness has both intensive and extensive margin components. A continuing establishment with a positive innovation to profitability grows less and the relationship between productivity and survival has weakened.² This decline in the responsiveness of establishments to shocks can have quantitatively important implications for productivity. If establishments are less responsive to changes in fundamentals, then the reallocation of factors across plants may be impeded, thus reducing aggregate productivity.³

Building on this evidence, this paper studies four potential sources of the reduction in responsiveness. First, an increase in the costs of factor adjustment will naturally lead, through the plant-level policy function, to a muted response to shocks. Second, changes in the stochastic process can induce establishments to be less responsive to innovations to profitability.⁴ In the presence of adjustment costs, if shocks are less serially correlated, then factor demand will react less to innovations. Third, as factor demand is a forward looking decision, a change in the discount factor will impact the response to shocks.⁵ Finally, if establishments have increased market power, reflected through a change in the curvature of the revenue function, then the marginal return to adjusting factor inputs is reduced and thus so is responsiveness.

Our contribution to this literature is to estimate a structural dynamic labor choice model at the plant-level to identify and quantify the contribution of these alternative sources of the reduction in dynamism. In the model, variations in profitability induce job growth and destruction as well as exit. The magnitude of these responses is influenced by the costs of labor adjustment, the impatience of the decision-maker, market power and the stochastic process for the underlying profitability shocks.

The dynamic optimization problem is brought to the data through simulated method of moments estimation. The measure of responsiveness comes from regression coefficients linking employment growth at the plant-level to these shocks and is among the moments matched. The estimation is undertaken for two sample periods: the (i)

¹These statistics are drawn from the public domain Business Dynamic Statistics (<https://www.census.gov/programs-surveys/bds/data.html>) and reflect the average annual rates from 1980-89 and 2000-2019 respectively. The job reallocation rate is the sum of job creation and destruction rates. Our micro data analysis only focuses on the 1980-2010 period but the decline in dynamism persists post 2010. The job reallocation rate over the 2001-10 period averages 29.7 percent.

²Key citations underlying this evidence and discussion include Davis, Haltiwanger, Jarmin, and Miranda (2007), Decker, Haltiwanger, Jarmin, and Miranda (2014), Decker, Haltiwanger, Jarmin, and Miranda (2016) and Decker, Haltiwanger, Jarmin, and Miranda (2020).

³Another important component of the decline is accounted for by the decline in business startups and an accompanying shift in activity towards older firms. The decline in startups has been the focus of much research (see, e.g., Decker, Haltiwanger, Jarmin, and Miranda (2014), Hopenhayn, Neira, and Singhania (2018), and Karahan, Pugsley, and Sahin (2018)). While the decline in startups is of considerable interest, the evidence from Decker, Haltiwanger, Jarmin, and Miranda (2020) suggests that the changing age distribution of firms accounts for less than 30 percent of the overall decline in the pace of job reallocation. We abstract from the declining pace of startups in our analysis.

⁴As formalized in our model, profitability shocks impact profits and reflect both variations in productivity and demand.

⁵The effect is somewhat model specific depending in part on whether there is time to build.

1980s and (ii) the 2000s.⁶ The analysis focuses on manufacturing plants.

A main finding of the paper is that the reduction in responsiveness is largely a consequence of increased costs of adjusting labor. This is shown through a series of counter-factual exercises. In one, adjustment costs are kept at their estimated values from 1980, with other parameters set at their estimated 2000 level. In this case, the fit of the model is reduced dramatically indicating the importance of changes in adjustment costs over time as an explanation for the reduced responsiveness. There are also estimated changes in impatience and the stochastic process over the two decades. But, for the stochastic process, these changes led to an increase rather than a dampening of responsiveness. For the discount factor, when it is re-estimated to match the 2000 responsiveness, leaving other parameters at their 1980 values, its change alone is not sufficient to generate the observed reduction in responsiveness. Finally, variations in market power, seen through the curvature of the revenue function, are also not able to generate the observed reduction in responsiveness.

A change in responsiveness has consequences for reallocation and thus for aggregate productivity. Based upon these estimates, the increase in labor adjustment costs implies that the growth in aggregate productivity in the manufacturing sector is about 8 percentage points lower in the 2000s than it would have been in the absence of the increase. Additional related consequences of the increase in adjustment costs is an increase in the dispersion in revenue labor productivity across plants and a decline in the covariance between the plant-level revenue labor productivity and the employment labor share.

The increase in labor productivity dispersion has additional implications related to evidence of an increase in revenue-weighted markup of firms. The underlying methodology in this literature, as seen in De Loecker, Eeckhout, and Unger (2020), measures the markup as the ratio of the output elasticity of a variable factor to the cost share of revenue of that factor. From De Loecker, Eeckhout, and Unger (2020), the finding of an increase in the revenue-weighted mean markup is largely accounted for by dispersion in markups across firms. That is, markups are increasing in the revenue of the firm along with rising concentration of revenue. Our findings highlight the effects of adjustment costs for labor, that we estimate are increasing over time, on measured market power. Using the De Loecker, Eeckhout, and Unger (2020) methodology, which assumes labor is adjusted without frictions, our model generates substantial and rising dispersion in measured markups without any variation in actual markups. Moreover, matching data patterns, the measured markup is increasing in productivity and the revenue share of the business in our simulated data.

The paper proceeds as following. Section 2 provides motivating evidence of declining responsiveness closely linked to the recent literature. The model is specified in section 3. Estimation of the structural model via the method of simulated moments is presented in section 4. The data moments include the responsiveness moments from section 2 as well as estimates of the revenue function curvature and the standard deviation and persistence of the shock processes. Based on this estimation, we evaluate explanations and implications of the decline in dynamism in section 5. Concluding comments are in section 6.

⁶In our empirical analysis, the moments from the 1980s reflect the 1990-89 period and 2000s reflect the 2000-2010 period. We often refer to the moments in the 1980s as the 1980 moments and the 2000s as the 2000 moments.

2 Motivating Evidence

A starting point for our analysis is that the decline in measures of business dynamism (e.g., the pace of job reallocation) have been accompanied by declines in measured responsiveness. As discussed, Decker, Haltiwanger, Jarmin, and Miranda (2020) present evidence that the marginal response of establishment and firm-level employment growth to realizations of measured productivity have declined. The evidence most relevant for our analysis is for U.S. manufacturing establishments. They find a decline in responsiveness of manufacturing establishments to productivity shocks that includes both an intensive (lower responsiveness for continuing establishments) and extensive (the marginal effect of productivity on exit declines) component. Using an accounting exercise, they find that lower responsiveness accounts for virtually all of the measured decline in job reallocation. In contrast, they find that the dispersion of measured productivity across establishments in the same industry is rising. Digging further, the persistence of measured productivity shocks has changed little but the innovations to measured productivity within industries has increased. This work is primarily an empirical exploration of different alternative measures of productivity and variation in declining of responsiveness across different types of firms and establishments. The mechanisms underlying the decline in reallocation, the decline in responsiveness and the increase in measured productivity dispersion are not investigated.

In related work, Ilut, Kehrig, and Schneider (2018) also estimate the response of job growth to productivity innovations. Their focus is on the concavity of this relationship, which they verify using plant-level data, and its implications for the cyclicity of volatility. Building on this, Kehrig and Vincent (2017) find that the concave relationship has changed over the decades in a manner consistent with the reduced responsiveness documented in Decker, Haltiwanger, Jarmin, and Miranda (2020).⁷ Additional research that implies that taking into account nonlinearities is likely important is found in Cairo (2013). She documents that accompanying the decline in dynamism is an increase in the fraction of employment at establishments that make no change in employment.

Another important and relevant line of research is De Loecker, Eeckhout, and Unger (2020). They argue that markups have risen dramatically in the US since the 1980s, particularly among the largest firms. Though their analysis does not focus on responsiveness *per se*, they do argue that the higher markups might explain the reduction in responsiveness that motivates our analysis.

With this discussion as background, Table 1 makes clear our perspective on the fall in responsiveness through recent decades. These results and those that follow rely on the establishment-level Annual Survey of Manufactures (ASM) data from 1980-2010 (essentially the same data as Decker, Haltiwanger, Jarmin, and Miranda (2018)) integrated with the Longitudinal Business Database (LBD).⁸ We use the ASM to generate measures of profitabil-

⁷Bartelsman, Lopez-Garcia, and Presidente (2019) study the effects of reallocation on productivity across 9 European countries. In doing so, they run responsiveness regressions similar to (1), adding in cyclical effects. The focus is not on the reduction in responsiveness across decades but rather a comparison across countries. They find both differences in responsiveness across countries and in response to aggregate fluctuations. They attribute these differences to variations in market power and employment protection across countries. In terms of the cyclical changes in responsiveness, they discuss adjustment costs, financial stress and the effects of a global reduction in trade flows.

⁸Decker, Haltiwanger, Jarmin, and Miranda (2018) is an earlier working paper version of Decker, Haltiwanger, Jarmin, and Miranda (2020) and uses the core micro data infrastructure integrating the ASM and the LBD from Foster, Grim, and Haltiwanger (2016). We use this same micro data infrastructure in our analysis along with the full LBD for some additional moments as described below.

ity shocks and we use the LBD to construct measures of establishment-level growth for continuing and exiting manufacturing establishments. We construct decade averages of key moments shown in Table 1. It is important to emphasize that all of the moments stem from the annual data from the ASM and the LBD. In the structural estimation below we take time aggregation issues into account.⁹

Table 1: Responsiveness Moments

Inact	xrat	ζ_1	ζ_2	ξ_1
1980s				
0.197	0.100	0.113	-0.054	-0.081
2000s				
0.243	0.083	0.064	-0.035	-0.059

The moments here are: Fraction of employment at inactive establishments = $0.025 > \frac{\Delta e}{e} > -0.025$, xrat = exit rate for establishments, (ζ_1, ζ_2) are linear and quadratic response of establishment employment growth to profitability shock innovation calculated as the average value by year for each decade from (1). ξ_1 is the response of plant-level exit to profitability shock calculated as the average value by year for each decade from (2).

The inaction and establishment exit rate moments are computed from the LBD manufacturing establishments.¹⁰ The intensive margin responsiveness moments are estimated coefficients from the following regression:

$$g_{it} = \alpha_0 + [\alpha_{1a} + \alpha_{1b}tr + \alpha_{1c}tr^2]\varepsilon_{it} + [\alpha_{2a} + \alpha_{2b}tr + \alpha_{2b}tr^2]\varepsilon_{it}^2 + \alpha_3 \log(emp_{i,t-1}) + \alpha_4 X_{it} + \eta_{it} \quad (1)$$

where g_{it} is the employment growth rate at establishment i in period t using the DHS measure given by $g_{it} = (emp_{i,t} - emp_{i,t-1}) / (0.5 * (emp_{i,t} + emp_{i,t-1}))$ for continuing establishments, ε_{it} is the innovation to the within industry-year (log) profitability (we describe that estimation below), tr is a time trend, emp_{it} is employment and X_{it} is a set of controls.¹¹ We include a quadratic on the innovation in this specification permitting responsiveness to vary with the magnitude of the innovation in shocks. As noted above, nonlinear response to innovations in a similar specification is found in Ilut, Kehrig, and Schneider (2018).¹² We permit the responsiveness coefficients to

⁹A distinguishing feature of our empirical analysis that builds on Decker, Haltiwanger, Jarmin, and Miranda (2018) is that we are using the LBD to measure employment growth outcomes and exit and not the ASM. Using the latter yields spurious exit from ASM panel rotation. Moreover, focusing only on the plants that are in consecutive years of the ASM to analyze the intensive margin is restrictive since this will be a non-representative set of plants. Our use of the LBD mitigates this issue which we further address by using inverse propensity score weights to take into account the probability in any given year that a manufacturing plant in the LBD is sampled in the ASM. See Decker, Haltiwanger, Jarmin, and Miranda (2018) for more details about the construction of the inverse propensity score weights.

¹⁰The moments from the LBD (including those in Table 2) are decade averages of annual statistics. The annual statistics have been adjusted for time varying cyclical in a manner similar to the cyclical adjustments in our responsiveness regression. While this does not matter much, the objective is for the differences in decade averages to reflect low frequency variation that is the focus of this paper.

¹¹We follow Foster, Grim, and Haltiwanger (2016), Decker, Haltiwanger, Jarmin, and Miranda (2018) and Decker, Haltiwanger, Jarmin, and Miranda (2020) in our specification of controls. Specifically, we include state and year effects, firm size controls, a state-level cyclical control and the interaction of the state-level cyclical control with the innovation to profitability. Our specification is similar to that in Decker, Haltiwanger, Jarmin, and Miranda (2018) and Decker, Haltiwanger, Jarmin, and Miranda (2020). Those papers provide extensive evidence of a decline in responsiveness of establishment-level growth rates and exit to alternative measures of productivity. Our measure of profitability is closest to what Decker, Haltiwanger, Jarmin, and Miranda (2020) refer to as TFPP. The magnitude of the declines in responsiveness we document is very similar to the findings in those papers. The primary reason we don't just the moments from those papers is that we also include a quadratic in the innovation to capture the type of nonlinearities emphasized in Kehrig and Vincent (2017).

¹²Looking at Germany, France, Italy and Spain, Cooper, Horn, and Indraccolo (2023) use moments that capture both the state dependent choices on the extensive (to adjust employment or not) as well as the intensive margins (job creation or destruction).

vary over time using a quadratic trend. We compute the estimated coefficient for each year using the quadratic trend and take the average value of the estimated coefficients for each year by decade, reported as (ζ_1, ζ_2) , in Table 1.¹³

For the innovations to measured profitability we first estimate the revenue function as a function of inputs capital, labor, materials and energy using the control function approach of Wooldridge (2009). This permits us to estimate the revenue elasticities for the inputs in a consistent manner. Using these revenue elasticities, we compute the revenue function residuals as profitability shocks. We specify an AR(1) process and compute the innovations to these residuals which yields the ε_{it} .¹⁴ Our approach of using the revenue function residuals as the profitability shock is consistent with Cooper and Haltiwanger (2006) and Decker, Haltiwanger, Jarmin, and Miranda (2020).¹⁵

As explained below, this estimate of the curvature of the revenue function and the inferred stochastic properties for profitability are used in the structural estimation below. Specifically, the estimated curvature is $\tilde{\alpha}$, and the serial correlation and standard deviation for this shock process are denoted by $(\tilde{\rho}, \tilde{\sigma})$.

The exit responsiveness equation is similar, given by:

$$exit_{it} = \beta_0 + [\beta_{1a} + \beta_{1b}tr + \beta_{1c}tr^2] \log(A_{i,t-1}) + \beta_2 \log(emp_{i,t-1}) + \beta_3 X_{i,t} + \mu_{it} \quad (2)$$

where $exit_{it}$ is a dummy variable equal to one of the establishment exits in period t , $A_{i,t-1}$ is the lagged profitability shock.¹⁶

The moments in Table 1 highlight that the decline in dynamism and responsiveness have a number of distinct features. We find that accompanying the decline in dynamism is a substantial increase in the fraction of employment at inactive establishments. Exit declines by almost 20 percent from the 1980s to the 2000s. Employment growth for continuing establishments is increasing but concave with respect to innovations in profitability shocks. This responsiveness declines from the 1980s to the 2000s.¹⁷ We also find that the relationship between exit and the realization of productivity has weakened over time.

¹³That is, the reported parameters (ζ_1, ζ_2) are the decade averages of the relevant α and β coefficients from equations (1) and (2) incorporating the quadratic trend by year. The responsiveness coefficient estimates are robust to using simple decade dummies instead of a quadratic trend. The responsiveness estimates are also robust to permitting the lagged employment estimated coefficients to change over time.

¹⁴In our analysis, we assume all factors of production other than labor are variable. This enables us to use these revenue elasticities to compute the curvature of the revenue function with respect to labor after substituting for the optimal variable other factors. The estimated curvature of the revenue function with respect to labor is 0.721.

¹⁵As noted, the revenue function residual we use is similar to the TFPP measure in Decker, Haltiwanger, Jarmin, and Miranda (2020) and Decker, Haltiwanger, Jarmin, and Miranda (2018). Ilut, Kehrig, and Schneider (2018) in contrast use a TFPR measure of profitability. In principle this distinction could be important since the revenue residual is under the assumptions of our model (see below) a measure of fundamentals while TFPR will reflect fundamentals and endogenous prices. However, Decker, Haltiwanger, Jarmin, and Miranda (2020) show that declining responsiveness is robust to using either a revenue function residual measure or TFPR. The revenue residual we use sweeps out industry by year effects in the same manner as Decker, Haltiwanger, Jarmin, and Miranda (2018).

¹⁶The ξ_1 in Table 1 is computed by taking the average value of the estimated coefficients for each year by decade.

¹⁷Both the linear and quadratic terms decline. As shown in Figure 1 over the range of range of innovations to shocks from negative 100 log points to positive log points that there is a decline in responsiveness. Also, if one compares Figure 10 of Ilut, Kehrig, and Schneider (2018) to our Figure 1 below for the 1980s and 2000s the patterns are quite similar.

3 Model: Dynamic Optimization and Response

This section accomplishes two goals. First, it states the formal optimization problem that forms the basis of the parameter estimation and ultimately our study of the causes of the decline in responsiveness. Second, drawing on the previous discussion, the model is used to illustrate the candidate explanations for the decline in responsiveness.

3.1 Plant-Level Optimization

The model of establishment labor demand builds on Cooper, Haltiwanger, and Willis (2015) and Cooper, Gong, and Yan (2015). Let the state of the plant be (A, e_{-1}) where A denotes current profitability and e_{-1} is the employment level from the previous period.

At the start of a period, the plant has an option to continue operating or exit. That choice is given by:

$$V(A, e_{-1}) = \max(V^c(A, e_{-1}), 0) \quad (3)$$

so that, by assumption, there is no cost associated with exit. There is a fixed cost of operating each period given by Γ . If the plant continues in operation, its value is given by

$$V^c(A, e_{-1}) = \max_e R(A, e) - \Gamma - e\omega - C(e, e_{-1}) + \beta E_{A'|A} V(A', e) \quad \forall (A, e_{-1}). \quad (4)$$

In this continuation problem, the controls are the number of workers, e .¹⁸ Note that by assumption there is no time to build: workers hired in the current period provide labor services immediately.

In (4), $R(A, e)$ is a revenue function. Other factors of production like hours, capital and materials have been optimized out leaving an expression of revenue as a function only of state and control variables. The variable A is interpreted as a shock to profitability as it encompasses both variations in total factor productivity and variations in product demand. Assume $R(A, e) = A(e)^\alpha$ so that α parameterizes the curvature of the revenue function.¹⁹

Finally, the adjustment cost function, $C(e, e_{-1})$, is given by

$$\begin{aligned} C(e, e_{-1}) &= \frac{\nu}{2} \left(\frac{e - e_{-1}}{e_{-1}} \right)^2 e_{-1} + [\gamma_P (e - e_{-1}) + F_p] I(e - e_{-1} > 0) \\ &\quad - [\gamma_M (e - e_{-1}) - F_m] I(e - e_{-1} < 0) \end{aligned} \quad (5)$$

for $e \neq e_{-1}$. There are no adjustment costs when there is no change in the number of workers.²⁰

This function includes multiple costs. One is the traditional quadratic adjustment cost, parameterized by γ . The next two are linear adjustment costs. Here γ_P is a linear hiring cost and γ_M is a linear firing cost. When γ_P and γ_M are different this implies kinked adjustment costs that will generate an inaction region. These may

¹⁸Given our moments are annual, we abstract from variation in hours per worker. Decker, Haltiwanger, Jarmin, and Miranda (2020) explore quarterly production hours and find limited changes in responsiveness of hours per worker.

¹⁹This is a common specification in the dynamic factor demand literature. It can be generated by a constant returns to scale technology combined with a constant elasticity of demand function where α captures both factor shares and the elasticity of demand.

²⁰Our data is not rich enough to allow us to match gross hires and fires at the plant-level.

be thought of as recruiting and severance costs respectively. Finally, there are fixed adjustment costs, (F_p, F_m) , which also depend on whether the plant is hiring or firing. In the estimation, we will distinguish the two cases of piece-wise linear and fixed costs.²¹

The resulting policy functions are given by: $Z_\Theta(A, e_{-1}) \in \{0, 1\}$ and $e = \phi_\Theta(A, e_{-1})$. Here $Z(\cdot)$ is the exit decision where $Z(A, e_{-1}) = 0$ means exit and $Z(A, e_{-1}) = 1$ denotes the continuation of the establishment. In the event of continuation, $e = \phi_\Theta(A, e_{-1})$ represent the state contingent choice over employees.

The exit of plants is offset by an exogenous entry process that maintains the population of plants. Entering plants do so with median employment and a draw from the profitability distribution. We also assume that the plants take the wage as given. In the estimation, the wage helps pin down the median size of plants.

Note that these policy functions depend on the underlying structural processes and parameters of the revenue, compensation and adjustment cost functions as well as the discount factor. All of these influences are captured by the vector of parameters, denoted Θ . Our goal is to estimate Θ through a simulated methods of moments approach for different sample periods to determine which elements of this vector are responsible for observed changes in responsiveness.

3.2 From Parameters to Responsiveness

Our results rest on the mapping from key model parameters to moments, including the responsiveness of job growth to profitability innovations and the extensive margin, both in terms of inaction in employment adjustment and in the continuation of operations. For a given vector of parameters, the plant-level dynamic optimization model is solved to produce decision rules. A panel is simulated and the regression models given in (6) and (7) are estimated to produce the responsiveness moments.

To capture the response on the intensive margin, consider a version of (1):

$$g_{it} = \zeta_0 + \zeta_1 \log(\varepsilon_{it}) + \zeta_2 \log(\varepsilon_{it})^2 + \zeta_3 \text{lemp}_{i,t-1} + \eta_{it}. \quad (6)$$

This is a linear quadratic empirical model linking job creation, g_{it} , to the (log) innovation to profitability, ε_{it} , $\text{lemp}_{i,t-1}$ is the log of lagged employment and η_{it} is the error term.²² On the extensive margin of employment adjustment, the fraction of observations in which job growth is sufficiently close to zero is recorded as “inaction”. Here, as in the estimation that follows, inaction is job growth less than 2.5% in absolute value.

Finally, also on the extensive margin, a simplified version of (2) links the exit decision to lagged profitability:

²¹Cooper, Haltiwanger, and Willis (2015) also estimates a model with an opportunity cost of adjustment, $\lambda R(A, e)$. That model was considered here too but did not match the 1980 moments as well as the other specifications.

²²This specification is simpler than (1) in a couple of ways. First, this specification does not have the additional controls X_{it} that control for factors that are not present in the model. Second, this specification should be considered as identifying responsiveness in the simulated data for a particular sub-period. It is also useful to note that the ε_{it} used here is based on using the curvature parameter estimated via the control function approach described in section 2. This curvature parameter is 0.721. By using this curvature parameter, we insure that the innovation process we use in the simulated data is consistent with what we measure in the actual data. Note that we still permit the curvature of the revenue function to be an estimated parameter in the simulated method of moments.

$$exit_{it} = \xi_0 + \xi_1 \log(A_{i,t-1}) + \xi_2 lemp_{i,t-1} + \mu_{it}. \quad (7)$$

The focus is on the coefficient on profitability, termed ξ_1 , as this captures the responsiveness of the exit decision to the log of lagged profitability, $A_{i,t-1}$.²³

From the perspective of these moments, a reduction in responsiveness means that job growth is less sensitive to shocks in (6), inaction is more frequent, and the exit decision is less responsive to profitability. The point of the empirical exercise is to determine which structural parameters can generate these forms of a reduction in responsiveness.

4 Estimation

The first step in the quantitative analysis is to estimate the parameters of the plant-level optimization problem. The estimation is conducted for two sample periods, the 1980s and the 2000s, to reflect the underlying theme of the reduced responsiveness of plants across these two decades. Once the parameter estimates are obtained, we will explore the factors that lead to the reduced responsiveness as well as the productivity implications.

4.1 Approach

The estimation relies on a simulated method of moments approach. The model parameters are selected to minimize the distance between actual and simulated moments, given by:

$$\mathcal{L} = \frac{(M^d - M^s(\Theta))'}{M^d} W \frac{(M^d - M^s(\Theta))}{M^d}. \quad (8)$$

In this expression, M^d are the data moments, $M^s(\Theta)$ are the simulated moments that, of course, depend on the parameters and W is a weighting matrix, here set as the identity matrix.

4.1.1 Moments

The full set of moments appear as the first row of Table 2. We bring together multiple moments that have been emphasized in different papers in order to facilitate identification of the different potential mechanisms at work.

For this analysis, the data moments include those from Table 1. Thus the responsiveness patterns that motivate the analysis are included as targets. In addition, we include the median plant size which has fallen slightly over the two sample periods. Including this moment insures that the employment state space in the model (roughly) conforms to that in the data. Among other things, this size is influenced by the wage, ω .

As in Cooper, Gong, and Yan (2015), the estimation includes the parameters of the revenue function as well as the driving processes for profitability. The alternative, as in Cooper and Haltiwanger (2006), is to infer these parameters from revenue alone. The problem with the latter approach is well known: omitted variable bias.

²³Since there is no data for a plant that exits in period t , there are necessarily lagged variables in this regression.

By including as moments the reduced form estimates of the revenue curvature and the stochastic process for the profitability shocks, the structural counterparts can be identified through the simulated method of moments approach.²⁴

Table 2: Moments

	Inact	xrat	ζ_1	ζ_2	ξ_1	emp	$\hat{\alpha}$	$\hat{\rho}$	$\hat{\sigma}$	\mathcal{L}
1980										
Data	0.197	0.100	0.113	-0.054	-0.081	10.100	0.977	0.687	0.368	na
Linear	0.201	0.053	0.149	-0.061	-0.140	10.064	0.937	0.394	0.336	1.050
Fixed	0.384	0.054	0.165	-0.056	-0.119	9.658	0.949	0.461	0.342	1.664
2000										
Data	0.243	0.083	0.064	-0.035	-0.059	8.900	0.959	0.682	0.408	na
Linear	0.214	0.053	0.065	-0.036	-0.108	8.759	0.918	0.350	0.369	1.089
Fixed	0.420	0.054	0.065	-0.033	-0.106	8.823	0.933	0.419	0.354	1.453

The moments here are: Inact = $0.025 > \frac{\Delta e}{e} > -0.025$ xrat = exit rate, (ζ_1, ζ_2) = linear and quadratic response of employment growth to profitability shock; ξ_1 = response of plant-level exit to profitability shock innovation; **emp** is median plant size. $(\hat{\alpha}, \hat{\rho}, \hat{\sigma})$ are the OLS estimates of revenue curvature as well as the serial correlation and standard deviation of the profitability shock. \mathcal{L} is the fit measured as percent deviation of simulated and data moments. All moments are from annual data.

Specifically, the curvature parameter in Table 2 is from an OLS regression of log revenue on the log labor input using annual data for each sample period. For our structural estimation, this OLS estimate, denoted $\hat{\alpha}$, is treated as a moment. The same regression is run on the simulated data. This is informative about the underlying curvature of the revenue function, α , in the vector of estimated parameters.

The moments also include the serial correlation and standard deviation of the innovation to the profitability shock process, denoted $(\hat{\rho}, \hat{\sigma})$ in Table 2. These data moments come from the control function approach described in section 2. Their simulated counterpart come from the simulated revenue and employment choices at the plant level along with the estimated revenue curvature parameter based upon the estimation of the revenue function in section 2. We are effectively treating the revenue curvature estimate from section 2 as a parameter that is used with the actual and simulated data to compute profitability shocks. This implies that the simulated and data moments are calculated in the same manner.

Interestingly, these estimates of the curvature of revenue and the underlying stochastic process might interact with the responsiveness moments. If, for example, the reduction in responsiveness is due to larger adjustment costs, then the omitted variable bias in the OLS regression of revenue on employment will be reduced as well. This interaction is fully taken into account in our methodology. The process for the shocks is also impacted here since it depends on the relationship between revenue and employment.

In using these moments, there are a few key points. First, though the moments are measured on an annual basis, we model the decision period as a quarter which we think of as closer to the frequency of choices made at the plant level. Thus matching the annual moments involves time aggregation of quarterly choices.

²⁴This step is needed also as we are studying an optimization problem at the quarterly frequency and do not have access to quarterly data on revenues and labor input.

Second, as made clear, the model includes exit in order to match various moments associated with plant closings. As discussed above, the exit is offset by an exogenous entry process.

Third, the data moments in Table 2 are for both the 1980 and 2000 time periods. Thus, we conduct a separate estimation for the 1980 and 2000 time periods.

4.1.2 Parameters

The parameter vector is given by: $\Theta = (\beta, \nu, \gamma_P, \gamma_M, f_P, f_M, \Gamma, \omega, \alpha, \rho, \sigma)$. All of these parameters were discussed in section 3. Here α is the structural parameter characterizing the curvature of the revenue function and (ρ, σ) characterize the stochastic profitability shock process. These are distinct from the OLS annual counterpart of curvature ($\hat{\alpha}$) and the parameters of the shock process computed from annual revenue and employment ($\hat{\rho}, \hat{\sigma}$) that are included in the moments.

There are two parameterizations of the problem. One assumes linear adjustment costs and the second assumes fixed adjustment costs. For each of the two parameterizations, there are 9 parameters and 9 moments. The model is just identified.

4.2 Estimates

Table 2 presents the moments (data and simulated) and Table 3 the parameters estimates. There are two sample periods and, for each, two specifications of adjustment costs.

From the last column of Table 2, the models with piece-wise linear adjustment costs match the data moments considerably better than do the models with fixed adjustment costs. This is true for both sample periods. We find that $\gamma_P < \gamma_M$ in both periods with the gap widening in the second period.

Looking first at the moments, the linear model matches quite well the inaction in labor adjustment while the fixed cost model misses this moment by a considerable margin. This is the main reason the piece-wise linear model fits better. Of course this does not mean the fixed cost model cannot create less inaction. But, to do so it would have to miss matching other moments. Both models capture the linear and quadratic coefficients from the responsiveness regressions, in both samples. And both overstate the responsiveness of exit to variations in profitability. The fixed cost model matches the OLS curvature moment as well as the stochastic process of the shocks a bit better than the piece-wise linear specification. Neither match the serial correlation very well. Interestingly, the OLS curvature estimates exceed the actual estimated curvature, see Table 3, indicating the presence of omitted variable bias in both models.

In terms of the parameters, the quarterly discount factor ranges between 0.9798 to 0.9871 across specifications and samples. This translates to a range of annual discount rates of 0.9216 to 0.9494 (with the highest in the linear case for the 1980s). These are much lower than the rate normally assumed in dynamic choice models. Both specifications exhibit relatively large quadratic adjustment costs, and firing costs are much larger than hiring costs, particularly in the linear cost case.²⁵ As discussed in Cooper, Gong, and Yan (2015), the relatively low discount

²⁵Cooper, Haltiwanger, and Willis (2015) did not estimate the discount factor and did not consider linear adjustment costs. Cooper,

factor aides in identifying hiring from firing costs.

Our main interest is in understanding differences across the samples. This is explored in detail in the following sections. Here we comment on some key moment and parameter differences.

There are, of course, changes in all moments across these time periods. As made clear in the motivation, the responsiveness is lower in the later sample. This is indeed captured by both models, both for continuers and those who exit. The higher inaction rate is also present though the increase in the linear model is not as large as in the data. Also, both models match the lower exit rate.

The estimated curvature of the revenue function is lower in 2000 relative to 1980. This may be consistent with the findings of an increase in market power reported in De Loecker, Eeckhout, and Unger (2020). However, as explored in sub-section 5.3 below, De Loecker, Eeckhout, and Unger (2020) measure market power in a manner quite differently than us. We utilize their approach to inferring market power and this allows us to compare our findings with theirs.

Table 3: Parameter Estimates

	β	ν	γ_P	γ_M	f_P	f_M	Γ	ω	α	ρ	σ
1980											
Linear	0.9871	4.3716	4.7281	6.7756	na	na	0.8652	0.1424	0.5399	0.8785	0.5887
Fixed	0.9811	5.6742	na	na	0.0003	3.9959	0.7412	0.2772	0.6030	0.8753	0.5436
2000											
Linear	0.9826	5.2215	4.7473	7.4113	na	na	0.8543	0.1257	0.5231	0.8564	0.6288
Fixed	0.9798	7.2333	na	na	0.0007	4.3685	0.7238	0.2355	0.5867	0.8626	0.5641

The parameters here are: β = discount factor, ν = quadratic adjustment cost, (γ_P, γ_M) =linear hiring and firing costs, (f_P, f_M) =are the fractions of revenue lost from fixed hiring and firing costs, Γ = fixed production cost as a fraction of average revenue, ω =base wage, (α, ρ, σ) =curvature of revenue functions, serial correlation of profitability shocks and the standard deviation of the innovation to profitability shocks.

The top row of Figure 1 isolates the intensive responses to profitability for the two decades in both the data (left panel) and the estimated models (right panel). The reduction in responsiveness is clear in both the data and the models, as is the curvature in these responses. The bottom panels of this figure are used in the decomposition exercise, as explained below.

5 Explanations and Implications of the Decline in Dynamism

Clearly the estimated model picks up the reduction in responsiveness. This leads to two questions: (i) what is driving this reduction in responsiveness and (ii) what are the consequences for productivity?

Pursuing the first question, this section reports two exercises intended to focus specifically on the responsiveness moments. One is through a counter-factual exercise in which the four factors emphasized earlier are investigated individually. The second is through another estimation exercise intended to match just the change in responsiveness,

Gong, and Yan (2015) estimated an annual discount factor for private plants of 0.929 and found that firing costs exceeded hiring costs.

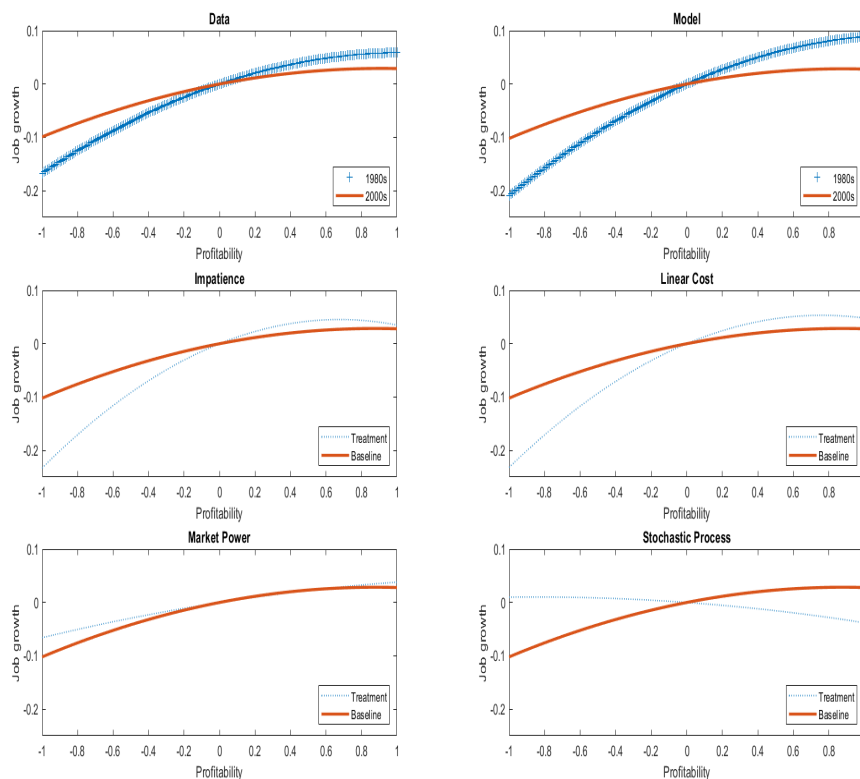


Figure 1: Job Growth Response To Innovations: Data and Model

For the top row, the left (right) panel is based upon coefficients from the responsiveness regression on actual (simulated) data, the latter from the model with linear adjustment costs. The other rows correspond to the experiments in Table 4.

using the four candidates. Taken together, the results of these two exercises lead to the conclusion that increased costs of labor adjustment are the primary contributor to the reduction in responsiveness. This does not imply that other parameter changes, such as the estimated reduction in patience, are immaterial. Rather, the preponderance of evidence points to higher adjustment costs as the main source of the reduced responsiveness over the decades.

To address the second question, we study the effects of the change in adjustment costs on a number of moments of the productivity distribution. The increased adjustment costs, by reducing responsiveness, yields an increase in the dispersion of revenue labor productivity across plants, a decline in the covariance between revenue labor productivity and the employment share of a business and a decline in aggregate productivity.

5.1 Sources

This section decomposes the changes in parameter estimates between 1980 and 2000 to detect the relative importance of the four explanations (adjustment costs, impatience, market power and expectations) for the reduction

in responsiveness. There are two approaches taken. The first is simulation based. For this exercise, we take a subset of parameters and set them at their 1980 estimates, allowing other parameters to remain at their estimated values for the 2000 sample moments. We use this to determine which parameter changes mattered most across the two samples in terms of matching both the responsiveness moments and all other moments. The second involves re-estimation of key parameters. In this case, a subset of the parameters are re-estimated to fit the 2000 responsiveness moments, holding the others fixed at their estimated values for 1980. The model fit is evaluated both in terms of matching the targeted moments as well as the untargeted ones.

From the combined evidence from these exercises, we argue that the main source of the reduction in responsiveness is the increase in labor adjustment costs. Within these costs, increases in all components, both quadratic and fixed, contribute to the fall in responsiveness.

What might be the sources of increases in adjustment costs? Decker, Haltiwanger, Jarmin, and Miranda (2020) and Davis and Haltiwanger (2014) suggest there are numerous sources of increased frictions in the adjustment of employment in the U.S. They provide evidence of changes in employment-at-will doctrines in the U.S. judicial system, rising prevalence of occupational licensing, increasing use of non-compete clauses, and potential indirect factors (such as zoning) that impair geographic labor mobility. We leave exploring the connection between these potential sources of increasing adjustment costs and our findings for future research. Our findings provide guidance and discipline on the nature and magnitude of the increase in adjustment costs that such factors must account for. Specifically, matching the decline in responsiveness requires increasing both convex and non-convex components of adjustment costs. Also, we think that the adjustment costs we identify may reflect broader costs of adjusting the scale of operations of a plant. For example, Decker, Haltiwanger, Jarmin, and Miranda (2020) find that the responsiveness of investment declines from the 1980s to the 2000s.

5.1.1 Simulation Based Decomposition

A simulation based decomposition of the changes in parameter estimates provides evidence that two of the explanations, changes in market power or the stochastic process, are not primary contributors to the reduction in responsiveness. Table 4 reports the results for the simulation based decomposition, where the top panel reports the data moments and the simulated moments from the best-fitting model with piece-wise linear adjustment costs, all for 2000. The second and third rows of Figure 1 illustrate responsiveness for these experiments. The dark curve is the 2000 baseline and the “treatment” comes from holding a particular set of parameters at their 1980 values.

Within the middle panel, each row corresponds to one of the four leading explanations for the reduction in responsiveness. For each of these, the associated parameters are kept at their estimated 1980 values, else parameters are at their 2000 estimated values. So, for example, in the third row of the second block, α is set at its 1980 estimated value of 0.5399 rather than its 2000 estimated value of 0.5231. With all other parameters at their estimated 2000 values, we evaluate the effect on the moments and therefore the fit of the small change in α . This exercise is repeated for the other cases.

From the last column, it is clear that changes in the stochastic process, followed closely by the discount factor

Table 4: Simulated Moments: Sources of Changes

	Inact	xrat	ζ_1	ζ_2	ξ_1	emp	$\hat{\alpha}$	$\hat{\rho}$	$\hat{\sigma}$	\mathcal{L}
2000 Data	0.243	0.083	0.064	-0.035	-0.059	8.900	0.959	0.682	0.408	na
1980 Baseline	0.201	0.053	0.149	-0.061	-0.140	10.064	0.937	0.394	0.336	1.050
2000 Baseline	0.214	0.053	0.065	-0.036	-0.108	8.759	0.918	0.350	0.369	1.089
β	0.212	0.040	0.134	-0.099	-0.093	8.295	0.975	0.381	0.379	5.334
$C(\cdot)$	0.202	0.048	0.140	-0.092	-0.109	8.435	0.948	0.363	0.369	5.262
α	0.219	0.056	0.052	-0.014	-0.109	14.630	0.914	0.356	0.373	1.894
(ρ, σ)	0.250	0.067	-0.024	-0.014	-0.158	9.902	0.863	0.343	0.333	5.420
$\beta, C(\cdot)$	0.187	0.035	0.212	-0.138	-0.085	8.123	1.007	0.404	0.380	14.667
(β, ρ, σ)	0.252	0.056	0.037	-0.065	-0.144	9.729	0.905	0.367	0.341	3.327
$(C(\cdot), \rho, \sigma)$	0.236	0.063	0.050	-0.051	-0.160	9.559	0.889	0.354	0.332	3.521

The moments here are: $\text{Inact} = 0.025 > \frac{\Delta e}{e} > -0.025$ xrat = exit rate, (ζ_1, ζ_2) = linear and quadratic response of employment growth to innovation to profitability shock; ξ_1 = response of plant-level exit to profitability shock; **emp** is median plant size. $(\hat{\alpha}, \hat{\rho}, \hat{\sigma})$ are the OLS estimates of revenue curvature as well as the serial correlation and standard deviation of the profitability shock. All moments are from annual data.

and the costs of adjustment, generate the largest losses in model fit. The fit rises from 1.089 in the baseline to over 5.2 for these cases. In contrast, the change in the estimated value of α is relatively unimportant.

The second and third rows of Figure 1 illustrate responsiveness for these experiments. The dark curve is the 2000 baseline and the “treatment” comes from holding a particular set of parameters at their 1980 values. These experiments suggest that the main difference across the decades came from either changes in impatience or in adjustment costs. From the bottom right panel, the change in the stochastic process actually leads to an increase in responsiveness, contrary to data patterns.

From the middle row of the figure, it is clear that setting the discount factor or the adjustment costs back to their 1980 values will increase responsiveness. The discount factor is lower in the 2000s and choices are quite sensitive to this parameter. Keeping β at its 1980 value increases the linear part of the responsiveness regression almost back to the 1980 level and increases (in absolute value) the coefficient on the quadratic term. The higher β also increases the OLS estimate of the curvature as the higher responsiveness implies more omitted variable bias. The estimated adjustment costs, particularly the quadratic and firing costs, rise over the two samples. From Table 4, the deterioration in fit in the treatment comes in large part due to the responsiveness moments. At the 1980 values of the adjustment costs, the two regression coefficients, (ζ_1, ζ_2) , are much higher than their values in the estimated model for 2000 and in the data. These results are similar to those from the experiment with β .

The difference between the treatment and baseline is very small in the market power experiment and when the shock process is set back to its 1980 values, responsiveness actually decreases. The estimate of the revenue curvature in 2000 is only slightly lower than that from 1980. From the bottom left part of Figure 1, there is very little difference between the baseline and treatment in terms of responsiveness. From Table 3, ρ is lower in the latter period. The reduction in the serial correlation will naturally reduce the responsiveness since the expected gain from employment adjustment is a dependent on future values of profitability. However, σ is higher and this

increases responsiveness. The lower right panel of Figure 1 shows that changing the stochastic processes actually goes the wrong way in terms of responsiveness relative to the data (top panel). That is, based on the stochastic processes alone the responsiveness would be lower in the 1980s than the 2000s. This poor performance is evident in the estimated coefficients. For example, ζ_1 changes sign relative to the best fit case.

The bottom panel of the table considers changes in pairs of parameters. Holding both the discount factor and the adjustment costs at their 1980 estimated values, the fit of the model deteriorates further, almost tripling the fit measure (14.667) compared to the first set of experiments. This is not the case when the stochastic process is coupled with either the discount factor or adjustment costs. For these two cases, the fit does not deteriorate as much, in part because of the offsetting effects on the responsiveness moments.

These experiments, summarized by both the moments and Figure 1, suggest that the main contributors to the decreased responsiveness across the decades came from either changes in impatience or in adjustment costs. From the bottom right panel of the figure, the change in the stochastic process actually leads to an increase in responsiveness, contrary to data patterns. The analysis in the next subsection provides additional evidence to help distinguish the main source of the reduction in responsiveness

5.1.2 Reestimation Based Decomposition

A reestimation based decomposition provides additional evidence that points to adjustment costs as the primary factor contributing to the reduction in responsiveness. This alternative perspective on the source of the reduction in responsiveness comes from an estimation exercise in which most parameters are kept at their 1980s values, allowing only subsets to be re-estimated. Further, the re-estimation itself focuses solely on the three moments capturing responsiveness: the two regression coefficients from the job growth regression and the response of exit to profitability.²⁶ Thus the exercises addresses the question if any of the four mechanisms we study could explain the reduction in responsiveness alone, leaving aside the other moments.

Table 5: Targeted Moments: Sources of Changes

Case	Targeted			\mathcal{L}_{targ}
	ζ_1	ζ_2	ξ_1	
Data 1980	0.113	-0.054	-0.081	na
Data 2000	0.064	-0.035	-0.059	na
Baseline 2000	0.065	-0.036	-0.108	0.691
β	0.097	-0.035	-0.164	3.433
$C()$	0.055	-0.035	-0.145	2.145
α	0.140	-0.088	-0.144	5.779
(ρ, σ)	0.134	-0.031	-0.123	2.402

The moments here are: (ζ_1, ζ_2) = linear and quadratic response of employment growth to innovation to profitability shock; ξ_1 = response of plant-level exit to profitability shock. \mathcal{L}_{targ} measures the fit of the model to these moments alone. All moments are from annual data.

²⁶For this exercise, we term these targeted moments.

For this exercise, Table 5 reports the targeted and data moments while Table 6 presents the estimates. The rows correspond to the leading explanations for the reduction in responsiveness.

Looking at the $C(\cdot)$ row, the estimates for the three adjustment costs are given in Table 6. Relative to their 1980 and 2000 estimated values, both the linear firing and hiring costs are higher but the quadratic cost is actually lower. Of course, the estimation exercises differ both in terms of the moments matched and in fixing other parameters at their 1980 baseline. From Table 5 the 2000 responsiveness moments are matched quite well and the fit is the closest among these experiments. Of course, there are also three parameters that are being reestimated.

As for the impatience experiment, from Table 6 a value of $\beta = 0.9815$, slightly lower than the original 2000 estimates. This lower value of the discount factor reduces the responsiveness as shown in Table 5. This reduction in β creates large responses to profitability on the intensive and on the exit margins compared to the 2000 data response and so the fit is not nearly as good as the adjustment cost case. Thus, this second exercise rules out impatience as the primary factor contributing to a reduction in responsiveness.

For the stochastic process, the serial correlation is about at its 2000 baseline estimate but the re-estimated value of σ is lower. As with the impatience experiment, the linear responses on both the intensive and extensive margins exceed those of the data. In fact, for this case the linear response on the intensive margin exceeds that of the 1980 data.

As in the previous exercise, variations in α do not produce the reduction in responsiveness seen in the data. The point estimate of α is slightly lower than the baseline 2000 estimate but the responsiveness is not close to the data.

Table 6: Re-estimated Parameters: Sources of Changes

	β	ν	γ_P	γ_M	Γ	ω	α	ρ	σ
Baseline 1980	0.9871	4.3716	4.7281	6.7756	0.8652	0.1424	0.5399	0.8785	0.5887
Baseline 2000	0.9826	5.2215	4.7473	7.4113	0.8543	0.1257	0.5231	0.8564	0.6288
β	0.9815	4.3716	4.7281	6.7756	0.8652	0.1424	0.5399	0.8785	0.5887
$C(\cdot)$	0.9871	4.1702	4.7826	9.1077	0.8652	0.1424	0.5399	0.8785	0.5887
α	0.9871	4.3716	4.7281	6.7756	0.8652	0.1424	0.5222	0.8785	0.5887
(ρ, σ)	0.9871	4.3716	4.7281	6.7756	0.8652	0.1424	0.5399	0.8565	0.6009

5.2 Productivity

Here we look at the effects on aggregate productivity of the changes in parameters isolated in the previous discussion. We do so at the quarterly frequency to highlight productivity implications absent time aggregation.²⁷

Table 7 provides insights into the productivity implications of the increase in adjustment costs. The productivity measures here are (i) the time series average of aggregate revenue per worker, a measure of aggregate productivity, (AggProd), (ii) the time series mean of the cross sectional standard deviation of the average revenue product

²⁷Interestingly, the productivity implications are influenced by time aggregation. For example, in the no adjustment cost baseline case, the correlation between the average revenue product of labor and the employment share is, as it should be, near 0 in the quarterly data. But time aggregation generates a negative correlation on the annual basis.

of labor (Mstd), (iii) the time series mean of the cross sectional covariance between the profitability shock and employment, $c(A, e)$. These last two statistics capture the productivity implications of the misallocation of labor.

Table 7: Productivity Implications

Sample	AggProd	Mstd	$corr(A, e)$
1980	1	7.147	0.768
2000	0.925	7.868	0.720

The statistics are computed from simulated data with best fit parameters from estimation. Frequency is quarterly. AggProd=1 in 1980 as a normalization.

Table 7 indicates three dimensions of productivity created from simulating our estimated models for the two sample periods. From our measure of AggProd, the factors that contributed to the reduction in responsiveness led, all else the same, to a reduction in productivity of about 8%. The reduction in reallocation is seen through the increased dispersion in the average revenue product of labor, Mstd, and the lower correlation between profitability and employment. This translates into a reduction in aggregate productivity.

To put these numbers into context, the official statistics from the Bureau of Labor Statistics show an increase in U.S. manufacturing productivity of 29 percent from the 1980s to the 2000s.²⁸ Our results imply that without the estimated increase in labor adjustment costs it would have risen by nearly 37 percent. This implies a non-trivial drag on the increase in productivity from the reduction in responsiveness.

5.3 Markups

De Loecker, Eeckhout, and Unger (2020) argue that markups, particularly of large (revenue-based weights) firms, have risen considerably since the 1980s. They suggest that this may be the source of the observed reduction in responsiveness.

Our estimation results find only modest support for variations in the curvature of the revenue function, as an explanation for the reduction in responsiveness: our estimate of the curvature is slightly lower for the 2000s vs the 1980s sample. This is not necessarily inconsistent with the De Loecker, Eeckhout, and Unger (2020) findings in a couple of ways. First, the estimated curvature of the revenue function reflects both markups and factor shares. But, importantly, it is the curvature of the revenue function that matters for responsiveness, not the markup *per se*. Second, they report that the unweighted average markup has not changed nearly as much as that of the revenue-weighted markup. Relatedly, they find that much of the increase in the revenue-weighted markup reflects reallocation towards large revenue firms that have higher markups. Our analysis is thus far silent about the distribution of the markup across heterogeneous firms.

There is another important difference worth pursuing: our model has explicit labor adjustment costs while their

²⁸For this purpose we use the increase in total factor productivity from the BLS for the US manufacturing sector. We use the growth in TFP from the average of the 1980-89 to the average of the 2000-10 period (equivalent to the time period of our micro data infrastructure). Our model environment only has one factor so this equivalent to total factor productivity in our model setting. We have not considered the possible increase in capital adjustment costs over this period.

estimation relies on the presence of a variable input without such costs.²⁹ This opens up intriguing questions. First, what distortions in the measurement of markups, using their approach, are created by the presence of adjustment costs? Can the model generate cross-sectional variation in the markup even if the curvature of the revenue function is constant across plants? Second, could changes in these adjustment costs account for their findings of increased markups comparing the 1980s to the 2000s? We answer these questions by following the De Loecker, Eeckhout, and Unger (2020) approach to inferring markups from revenue shares of variable inputs. The point is to understand what that approach would infer from simulated data about time series variation in the markup created by the changes across decades in our estimated model with labor adjustment costs.

Specifically, De Loecker, Eeckhout, and Unger (2020) calculate the markup, defined as the ratio of price over marginal cost, through a first-order condition from cost minimization, without adjustment costs. This yields:

$$\mu_{it} = \theta_{it}/ls_{it} \tag{9}$$

where μ_{it} is the markup, θ_{it} is the output elasticity of labor and ls_{it} is the share of **total revenue** that is paid to labor.³⁰ Throughout, i is a plant and t is time. This measurement approach for estimating markups is often denoted the production or ratio approach.

In the absence of adjustment costs and with fixed factor and demand elasticities, the labor share will be equalized across firms. However, adjustment costs yield variation in the labor share across firms even in the presence of fixed factor and demand elasticities. To explore how much variation we obtain in measured markups using the adjustment costs, we proceed as follows. From our simulated data, we can uncover ls_{it} . We begin by setting $\theta_{i,1980} = 0.673$ for all i , which is the value required to match the revenue-weighted mean markup in De Loecker, Eeckhout, and Unger (2020) in the early 1980s (from Figure 6 of their paper). We fix this value and then simulate the model using the estimated parameters for the 1980 and 2000 decades (including the estimated variation in α).

Results are reported in Table 8. We find that measured markups increase non-trivially between the 1980s and 2000s – about half of the magnitude of the increase in markups in US manufacturing over this period reported by De Loecker, Eeckhout, and Unger (2020). We also compute a number of additional moments of markups. Specifically, we compute the revenue-weighted median, revenue-weighted 90th percentile, the correlation of the markup with the market share in terms of revenue, and the correlation of the markup with productivity. The first three of these moments are reported in De Loecker, Eeckhout, and Unger (2020) and they also report results that imply a positive relationship between markups and market share as well as productivity.

Our simulated model with adjustment costs yields substantial dispersion in measured markups and a positive relationship between measured markups and market share as well as a positive relationship between measured markups and productivity. This is true for the cross sections in both 1980 and 2000. This is in spite of there being no cross-sectional variation in actual markups in our framework and only a modest reduction in α across the two

²⁹Some of their estimation uses combined labor and materials as the freely adjustable factor, other sections use just labor or just materials.

³⁰Bond, Hashemi, Kaplan, and Zoch (2021) argue that this revenue based measure of labor share implies, as a matter of theory, that the resulting markup is 1 regardless of the true markup when there are no labor adjustment costs.

sample periods. The increase in dispersion in markups is not as large as reported by De Loecker, Eeckhout, and Unger (2020) but there are likely other factors at work beyond adjustment costs that account for the increase in measured markup dispersion without increases in the actual markup dispersion.³¹

The bottom panel of Table 8 provides some counterfactuals intended to decompose the changes in the markup distribution over time by setting some parameters back to their 1980 values. From these results, there is no single mechanism that explains the findings. Holding adjustment costs at their 1980 estimates reduces the mean markup slightly. The lower estimate of α in 2000 plays a slightly larger role: setting this curvature at its 1980 value leads to a larger reduction in the average markup. Coupling these, the mean markup falls to 1.62. Combining this with setting the stochastic process at its 1980 value almost reproduces the 1980 findings.

Table 8: Moments of Measured Markups

	Mean μ	Median μ	P90 μ	$Corr(\mu, \frac{R}{\sum R})$	$Corr(\mu, A)$
	1980s				
Data	1.55	1.40	2.40	na	na
Model	1.55	1.50	2.12	0.18	0.45
	2000s				
Data	1.80	1.65	3.20	na	na
Model	1.69	1.61	2.44	0.20	0.48
set at 1980 est.	2000s Decompositions				
$C(\cdot)$	1.67	1.61	2.39	0.23	0.50
α	1.64	1.57	2.36	0.17	0.47
β	1.67	1.61	2.42	0.27	0.54
ρ, σ	1.64	1.58	2.30	0.13	0.43
$C(), \alpha$	1.62	1.56	2.32	0.21	0.49
$C(), \alpha, \rho, \sigma$	1.57	1.52	2.19	0.14	0.44

The markup measures follow De Loecker, Eeckhout, and Unger (2020) and are revenue weighted. Here P90 is the 90th percentile. The model moments are computed from simulated data with best fit parameters from estimation. Frequency is quarterly.

We recognize that we have not considered mechanisms that yield a true rise in revenue-weighted markups due to a combination of heterogeneous true markups and a shift in activity towards high markup firms (as in De Loecker, Eeckhout, and Mongey (2021)). However, the evidence of such a change in revenue-weighted markups due to reallocation effects to high markup firms is reliant on the indirect production approach. Our point here is that other mechanisms such as rising adjustment costs can yield an increase in the revenue-weighted dispersion, skewness and average of measured markups using the production approach.

³¹Foster, Haltiwanger, and Tuttle (2020) provide evidence that rising dispersion in factor elasticities accounts for a substantial fraction of rising dispersion in measured markups and the revenue weighted mean. Relatedly, Bond, Hashemi, Kaplan, and Zoch (2021) highlight the challenges of estimating factor elasticities using revenue and input data.

6 Conclusions

The point of this paper is to assess various explanations for the observed reduced responsiveness in labor demand to variations in profitable opportunities. The evidence is low frequency, a comparison between the 1980s and the 2000s. This reduction in responsiveness can have adverse aggregate productivity implications insofar as it reflects limitation to the process of factor reallocation.

Our approach uses simulated method of moments to estimate parameters of a plant-level optimization problem to match patterns across these two decades. The moments include the type of responsiveness measures that have sparked this literature.

Our main findings are easily summarized: much of the reduction in reallocation stems from increased costs in labor adjustment. While other explanations, such as changes in impatience, market power and the stochastic process governing revenues could also reduce responsiveness, these other variations were unable to do so and simultaneously match other moments. The increase in adjustment costs itself seems to be broad based, including both quadratic as well as the fixed costs of hiring and firing workers. Increased costs of labor adjustment are reflected in productivity losses.

Based on our estimates, aggregate productivity in U.S. manufacturing would have been 8 percentage points higher in the 2000s if adjustment costs had remained at their 1980 estimated levels. The increased frictions also imply increased dispersion in revenue labor productivity across businesses.

The increase in labor productivity dispersion has additional implications for measuring markups using the production approach, used for example in De Loecker, Eeckhout, and Unger (2020). Our findings imply substantial and rising dispersion in measured markups from the production approach without any variation in actual markups across firms. Moreover, we find that the measured markup at the micro level is increasing in productivity and the revenue share of the business. Looking across decades, the model generates about half of the increased markup reported in De Loecker, Eeckhout, and Unger (2020), with only a modest reduction in the curvature of the revenue function. Other factors (e.g., the type of superstar effects highlighted by Autor, Dorn, Katz, Patterson, and Van Reenan (2020)) that account for rising concentration can also yield an increase in the revenue-weighted mean markup in combination with the dispersion in measured markups from the adjustment costs.

Future research should extend this structural analysis in a number of directions. First, we permitted no structural heterogeneity across firms by observable characteristics such as industry, firm age and firm size. Given the observed structural changes even within manufacturing on these dimensions, such heterogeneity might be important for accounting for declining dynamism. Moreover, such structural heterogeneity may be important in accounting for higher moments of the productivity and in turn measured markup distribution. Second, this type of structural analysis should be extended beyond manufacturing. While it is more challenging to measure micro level profit shocks outside of manufacturing, the evidence shows that the declines in the pace of job reallocation is even more dramatic. The simulated method of moments approach we use has the potential to overcome these measurement limitations since it permits using the same restrictions in the simulated moments as in the actual data moments.

References

- AUTOR, D., D. DORN, L. F. KATZ, C. PATTERSON, AND J. VAN REENAN (2020): “The Fall of the Labor Share and the Rise of Superstar Firms,” Quarterly Journal of Economics, forthcoming.
- BARTELSMAN, E., P. LOPEZ-GARCIA, AND G. PRESIDENTE (2019): “Labour Reallocation in Recession and Recovery: Evidence for Europe,” National Institute Economic Review, 247(1), R32–R39.
- BOND, S., A. HASHEMI, G. KAPLAN, AND P. ZOCH (2021): “Some unpleasant markup arithmetic: Production function elasticities and their estimation from production data,” Journal of Monetary Economics, 121, 1–14.
- CAIRO, I. (2013): “The Slowdown in Business Employment Dynamics: The Role of Changing Skill Demands,” Mimeo.
- COOPER, R., G. GONG, AND P. YAN (2015): “Dynamic Labor Demand in China: Public and Private Objectives,” The RAND Journal of Economics, 46(3), 577–610.
- COOPER, R., AND J. HALTIWANGER (2006): “On the Nature of the Capital Adjustment Costs,” Review of Economic Studies, 73, 611–34.
- COOPER, R., J. HALTIWANGER, AND J. L. WILLIS (2015): “Dynamics of labor demand: Evidence from plant-level observations and aggregate implications,” Research in Economics, 69(1), 37–50.
- COOPER, R., C.-W. HORN, AND L. INDRACCOLO (2023): “Covid and Productivity in Europe: A Responsiveness Perspective,” *European Economic Review*, forthcoming.
- DAVIS, S. J., AND J. HALTIWANGER (2014): “Labor Market Fluidity and Economic Performance,” Federal Reserve Bank of Kansas City Jackson Hole Conference Symposium.
- DAVIS, S. J., J. HALTIWANGER, R. JARMIN, AND J. MIRANDA (2007): “Volatility and Dispersion in Business Growth Rates: Publicly Traded versus Privately Held Firms,” NBER Macroeconomics Annual 2006, 21, 107–80.
- DE LOECKER, J., J. EECKHOUT, AND S. MONGEY (2021): “Quantifying Market Power and Business Dynamism in the Macroeconomy,” NBER Working Paper No. 28761.
- DE LOECKER, J., J. EECKHOUT, AND G. UNGER (2020): “The rise of market power and the macroeconomic implications,” The Quarterly Journal of Economics, 135(2), 561–644.
- DECKER, R., J. HALTIWANGER, R. JARMIN, AND J. MIRANDA (2014): “The role of entrepreneurship in US job creation and economic dynamism,” Journal of Economic Perspectives, 28(3), 3–24.
- DECKER, R. A., J. HALTIWANGER, R. S. JARMIN, AND J. MIRANDA (2016): “Declining business dynamism: What we know and the way forward,” American Economic Review, 106(5), 203–07.

- (2020): “Changing Business Dynamism and Productivity: Shocks vs. Responsiveness,” American Economic Review, forthcoming.
- DECKER, R. A., J. C. HALTIWANGER, R. S. JARMIN, AND J. MIRANDA (2018): “Changing business dynamism and productivity: Shocks vs. responsiveness,” Discussion paper, National Bureau of Economic Research.
- FOSTER, L., C. GRIM, AND J. HALTIWANGER (2016): “Reallocation in the Great Recession: Cleansing or Not?,” Journal of Labor Economics, 34(S1), S293–S331.
- FOSTER, L., J. HALTIWANGER, AND C. TUTTLE (2020): “Rising Markups or Changing Technology?,” NBER Working Paper No. 30491.
- HOPENHAYN, H., J. NEIRA, AND R. SINGHANIA (2018): “From Population Growth to Firm Demographics: Implications for Concentration, Entrepreneurship and the Labor Share,” NBER Working Paper No. 25382.
- ILUT, C. L., M. KEHRIG, AND M. SCHNEIDER (2018): “Slow to Hire, Quick to Fire: Employment Dynamics with Asymmetric Responses to New,” Journal of Political Economy, 126(5), 2011–2071.
- KARAHAN, F., B. PUGSLEY, AND A. SAHIN (2018): “Demographic Origins of the Startup Deficit,” Mimeo.
- KEHRIG, M., AND N. VINCENT (2017): “Growing productivity without growing wages: The micro-level anatomy of the aggregate labor share decline,” Discussion Paper 244, Economic Research Initiatives at Duke (ERID) Working Paper.
- WOOLDRIDGE, J. M. (2009): “On Estimating Firm-Level Productivity Functions Using Proxy Variables to Control for Unobservables,” Economics Letters, 104(3), 112–114.