FEDERAL RESERVE BANK *of* ATLANTA WORKING PAPER SERIES

# Declining Responsiveness at the Establishment Level: Sources and Productivity Implications

Russell Cooper, John Haltiwanger, and Jonathan Willis

Working Paper 2024-3a
February 2024 (Revised January 2025)

**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 quantity 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 Matthias Kehrig, 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, the Technology and Declining Economic Dynamism in 2020, the NBER Summer Institute in 2024 and the EUI 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, Chinese Economic Research Institute, Liaoning University, and Department of Economics at the European University Institute, russellcoop@gmail.com; John Haltiwanger, Department of Economics, University of Maryland, haltiwan@econ.umd.edu; or Jonathan L. Willis, Research Department, Federal Reserve Bank of Atlanta, jon.willis@atl.frb.org.

# 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. A related 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 leading hypothesized 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. A reduction in the volatility of shocks or less serially correlated shocks can yield less responsiveness in the presence of adjustment costs. Third, as factor demand is a forward looking decision, a change in the discount factor will impact the response to shocks. Finally, if establishments have increased market power, reflected through a change in the curvature of the revenue function, then the

[1] These 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.

[2] Key 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).

[3] Another important component of the decline in 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.

marginal return to adjusting factor inputs is reduced and thus so is responsiveness. Put differently, an increase in the average markup yields a larger change in prices in response to an innovation in productivity relative to quantities.[4]

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 creation 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 decades: the (i) 1980s and (ii) the 2000s. 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 the 1980s period, with other parameters set at their estimated 2000s 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 2000s responsiveness, leaving other parameters at their 1980s 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 produc-

[4]For market power, we focus on the average markup as reflected in the curvature of the revenue function. With variable markups that are increasing in productivity and size, an increase in the dispersion and skewness of markups can yield a reduction in responsiveness as in De Loecker, Eeckhout, and Mongey (2021). While our modeling approach does not consider this approach, we show that evidence on rising dispersion and skewness in measured markups may be due to rising adjustment costs even without any dispersion in actual markups.

tivity. 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 are 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 the revenue-weighted measured markup of firms as found in De Loecker, Eeckhout, and Unger (2020). This literature measures the markup indirectly as the ratio of the output elasticity of a variable factor to the cost share of revenue of that factor. This indirect-measure approach is based on using first order conditions for variable factors of production that don't face adjustment costs. One of the factors De Loecker, Eeckhout, and Unger (2020) include as variable is labor. Our findings highlight the presence of adjustment costs of labor that are increasing over time, such as adjustment costs, will be reflected in indirectly measured markups. We show that, even in the absence of actual dispersion in markups, the rising adjustment costs we identify yield an increased revenue-weighted mean and dispersion in these indirectly measured markups. While we can only account for about half of the measured increase in markups with the rise in adjustment costs, our findings suggest that the connection that De Loecker, Eeckhout, and Mongey (2021) make between rising markups and declining responsiveness may reflect in part the rise in adjustment costs that reduces responsiveness and increases measured markups.

The paper proceeds as follows. 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 of the decline in dynamism in section 5. Productivity and markup implications are explored in section 6. Concluding comments are in section 7.

## 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.[5]

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 cyclicality 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 discussed above. Additional research that implies that taking into account nonlinearities is likely important is found in Cairo (2013). She documents

[5]Kehrig and Vincent (2021) also present evidence of declining responsiveness in the US. Bartelsman, 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. Cooper, Horn, and Indraccolo (2023) use a similar model and estimation approach to this paper to study variation in responsiveness across European countries with a focus on the impact of COVID policies. They highlight the importance of variation in adjustment costs in accounting for cross country variation.

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. De Loecker, Eeckhout, and Mongey (2021) explore the implications of rising markups for declining responsiveness. As we noted in the introduction and will discuss in detail below, the connection between declining responsiveness and rising *measured* markups may reflect the role of rising adjustment frictions.

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).[6] We use the ASM to generate measures of idiosyncratic profitability shocks and we use the LBD to construct measures of employment growth for continuing and exiting manufacturing establishments.[7] 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.

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

$$g_{it} = \zeta_0 + \zeta_{1t}\varepsilon_{it} + \zeta_{2t}\varepsilon_{it}^2 + \zeta_3 log(emp_{i,t-1}) + \zeta_4 X_{it} + \eta_{it} \tag{1}$$

where $g_{it}$ is the employment growth rate at establishment $i$ in period $t$ using the Davis, Haltiwanger, and Schuh (1998) measure given by $g_{it} = (emp_{i,t} - emp_{i,t-1})/(0.5 * (emp_{it} + emp_{i,t-1})$ for continuing establishments, $\varepsilon_{it}$ is the innovation to the within industry-year (log) profitability (we describe

[6]See the data appendix for more discussion of the data and measurement.

[7]We follow Decker, Haltiwanger, Jarmin, and Miranda (2018) in sweeping out industry by year effects from our profitability shocks.

Table 1: Responsiveness Moments

| Inact | xrat | $\zeta_1$ | $\zeta_2$ | $\xi_1$ |
|---|---|---|---|---|
| | | **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, $(\zeta_1, \zeta_2)$ 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). $\xi_1$ is the response of plant-level exit to profitability shock calculated as the average value by year for each decade from (2).

that estimation below), $emp_{it}$ is employment and $X_{it}$ is a set of controls.[8] 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).[9]

We permit the responsiveness coefficients ($\zeta_{1t}$ and $\zeta_{2t}$) to vary across decades. To implement this we specify these coefficients as a quadratic in a time trend (e.g., $\zeta_{1t} = \beta_1 + \beta_{1b}tr + \beta_{1c}tr^2$ and the equivalent for $\zeta_{2t}$). We then 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 $(\zeta_1, \zeta_2)$, in Table 1.[10]

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, denoted as $A_{it}$. We specify an AR(1) process for profitability and compute the innovations to

[8] We 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.

[9] Looking 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).

[10] 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.

these residuals which yields the $\varepsilon_{it}$.[11] 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).[12]

The exit responsiveness equation is similar, given by:

$$exit_{it} = \xi_0 + \xi_{1t} log(A_{it-1}) + \xi_2 log(emp_{i,t-1}) + \xi_3 X_{i,t} + \mu_{it} \tag{2}$$

where $exit_{it}$ is a dummy variable equal to one of the establishment exits in period $t$, $A_{it-1}$ is the lagged profitability shock. We allow for $\xi_{1t}$ to vary across decades in the same manner as the responsiveness coefficients for the intensive margin[13] For the exit responsiveness equation we follow the literature that relates exit between the prior and the current period to the relevant state vector in the prior period (i.e., the lagged profitability shock and lagged employment).

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.[14] We also find that the relationship between exit and the realization of profitability has weakened over time.

As we have highlighted, the specifications we consider for the intensive margin of growth and

[11] In 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.

[12] As 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. We also note that Decker, Haltiwanger, Jarmin, and Miranda (2020) show declining responsiveness to the profitability shock $A_{it-1}$, the first difference of the profitablity shock and the innovation $\varepsilon_{it}$. We focus on the innovation based specification which is also the approach of Ilut, Kehrig, and Schneider (2018).

[13] The $\xi_1$ in Table 1 is computed by taking the average value of the estimated coefficients for each year by decade.

[14] Both 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.

the extensive margin of exit draw from the existing literature (e.g., Decker, Haltiwanger, Jarmin, and Miranda (2020) and Ilut, Kehrig, and Schneider (2018)). As we will see below these moments are informative about the structural parameters of the model we develop and estimate via the simulated method of moments.

# 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. We model the profitability shock process as an AR(1) in logs: $\log(A) = \rho \log(A_{-1}) + \varepsilon$ where the standard deviation of $\varepsilon$ is given by $\sigma$. The parameters $\rho$ and $\sigma$ are structural parameters of the stochastic process for profitability.

At the start of a period, where periods in the model are defined as quarters, the plant has an option to continue operating or exit. That choice is given by:

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

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

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

In the continuation problem, the control is the number of workers, $e$.[15] Note that by assumption

[15]Given our moments are annual, we abstract from variation in hours per worker. Decker, Haltiwanger, Jarmin,

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 $R(A, e) = Ae^{\alpha}$ so that $\alpha$ parameterizes the curvature of the revenue function.[16]

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

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

for $e \neq e_{-1}$. There are no adjustment costs when there is no change in the number of workers.[17]

This function includes multiple costs. One is the traditional quadratic adjustment cost, parameterized by $\gamma$. The next two are linear adjustment costs. Here $\gamma_P$ is a linear hiring cost and $\gamma_M$ is a linear firing cost. When $\gamma_P$ and $\gamma_M$ are different this implies kinked adjustment costs that will generate an inaction region. These may 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.[18]

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

and Miranda (2020) explore annual production worker hours per worker and find limited changes in responsiveness of hours per worker.

[16]This 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 $\alpha$ captures both factor shares and the elasticity of demand.

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

[18]Cooper, 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 1980s moments as well as the other specifications.

plants.[19] Entering plants do so with median employment and a draw from the profitability distribution. We also assume that the plants take the wage ($\omega$) 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 $\Theta$. Our goal is to estimate $\Theta$ through a simulated method 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 employment 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 versions of 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):

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

This is a linear-quadratic empirical model linking *annual* net employment growth, $g_{it}$, to the (log) innovation to profitability, $\varepsilon_{it}$; $lemp_{i,t-1}$ is the log of lagged employment and $\eta_{it}$ is the error term. This specification is essentially the same as (1) in that is uses annualized measures of the dependent and independent variables (time aggregating the quarterly simulated data).

The specification differs from the actual data specification that address issues in the actual data not present in the model. 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.[20] On

[19]This assumption limits the potential for changes in adjustments costs to impact the distribution of entrants.

[20]It is also useful to note that the $\varepsilon_{it}$ 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

the extensive margin of employment adjustment, the fraction of observations in which employment growth is sufficiently close to zero is recorded as "inaction." Here, as in the estimation that follows, inaction is employment growth less than 2.5 percent in absolute value.

Finally, also on the extensive margin, a simplified version of (2) (again omitting controls relevant in the actual data but not present in the model) links the exit decision to lagged profitability:

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

The focus is on the coefficient on profitability, termed $\xi_1$, as this captures the responsiveness of the exit decision to the log of lagged profitability, $A_{i,t-1}$.[21] Again, this specification is estimated at the annual frequency in the simulated data matching the approach with the actual data.

From the perspective of these moments, a reduction in responsiveness means that employment 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. Given our simulated method of moments approach, we use parsimonious specifications that yield moments that can be estimated in both the actual and simulated data that are informative for our structural model.

We interpret the parameters estimated from the actual data on responsiveness of growth and exit as equivalent to those estimated from the simulated data. The differences in specifications reflect controls in the actual data estimation for factors not present in the model.

# 4 Estimation

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

---

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.

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

## 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:

$$\pounds = \frac{(M^d - M^s(\Theta))'}{M^d} W \frac{(M^d - M^s(\Theta))}{M^d}. \tag{8}$$

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 identify matrix.

### 4.1.1 Moments

The full set of moments appear as the first row of Table 2. In addition to the responsiveness moments in Table 1, we bring together multiple moments that have been emphasized in different papers in order to facilitate identification of the different potential mechanisms at work. 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, $\omega$.

As in Cooper, Gong, and Yan (2015), the estimation includes the parameters of the revenue function as well as the driving processes for profitability. 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.[22]

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. It is **not** used as a direct estimate of the curvature of the revenue function. Instead it is use to infer the curvature of the structural model through an indirect inference exercise. That is, 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, $\alpha$, 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

[22]This 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.

Table 2: Moments

| | Inact | xrat | $\zeta_1$ | $\zeta_2$ | $\xi_1$ | emp | $\hat{\alpha}$ | $\hat{\rho}$ | $\hat{\sigma}$ | $\pounds$ |
|---|---|---|---|---|---|---|---|---|---|---|
| **1980s** | | | | | | | | | | |
| 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 |
| **2000s** | | | | | | | | | | |
| 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; $(\zeta_1, \zeta_2)$ = linear and quadratic response of employment growth to profitability shock; $\xi_1$= response of plant-level exit to profitability shock innovation; **emp** is median plant size. $(\hat{\alpha}, \hat{\rho}, \hat{\sigma})$ are the OLS estimate of revenue curvature as well as the serial correlation and standard deviation of the profitability shock. $\pounds$ is the fit measured as percent deviation of simulated and data moments. All moments are from annual data.

function approach described in section 2. Their simulated counterparts 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. We note that the $\varepsilon$ series used to compute the moments of the stochastic processes is also used in the responsiveness regressions. This implies that the simulated and data moments are calculated in the same manner on all dimensions.

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.

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 1980s and 2000s decades. Thus, we conduct a separate estimation for the two 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 $\alpha$ is the structural parameter characterizing the curvature of the revenue function and $(\rho, \sigma)$ characterize the stochastic (quarterly) 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. Importantly, this implies that differences between the $\hat{\rho}$ and $\hat{\sigma}$ moments and structural parameters $\rho$ and $\sigma$ in the estimation do not reflect a poor fit. A poor fit arises when the moments in the actual and simulated data differ.

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.[23]

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. Given this finding, most of our subsequent analysis focuses on the piece-wise linear adjustment cost specification.

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

[23] Our approach is to estimate the two types of adjustment costs individually to better understand the nature of adjustment costs in a sample as well as the changes over the periods.

Table 3: Parameter Estimates

| | $\beta$ | $\nu$ | $\gamma_P$ | $\gamma_M$ | $f_P$ | $f_M$ | $\Gamma$ | $\omega$ | $\alpha$ | $\rho$ | $\sigma$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **1980s** | | | | | | | | | | | |
| 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 |
| **2000s** | | | | | | | | | | | |
| 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: $\beta$ = discount factor, $\nu$= quadratic adjustment cost, $(\gamma_P, \gamma_M)$ =linear hiring and firing costs,$(F_P, F_M)$=are the fractions of revenue lost from fixed hiring and firing costs, $\Gamma$ = fixed production cost as a fraction of average revenue, $\omega$=base wage, $(\alpha, \rho, \sigma)$=curvature of revenue functions, serial correlation of profitability shocks and the standard deviation of the innovation to profitability shocks.

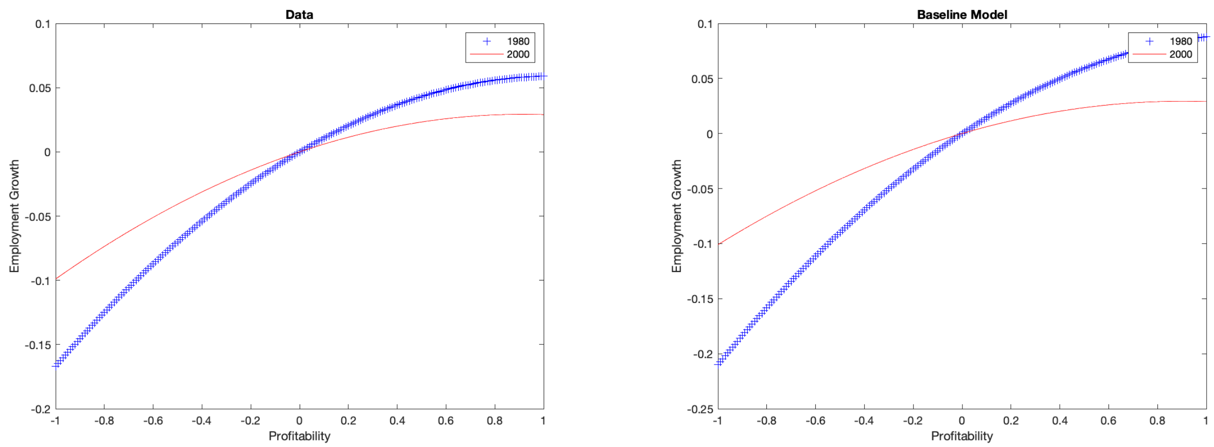
reason the piece-wise linear model fits better. Of course this does not mean the fixed cost model cannot create less inaction. But, to do so it would have to miss matching other moments. Both models capture the linear and quadratic coefficients from the responsiveness regressions, in both samples. And 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 matches 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.[24] As discussed in Cooper, Gong, and Yan (2015), the relatively low discount factor aides in identifying hiring from firing costs.

Our main interest is in understanding differences across the samples. This is explored in detail

[24] Cooper, Haltiwanger, and Willis (2015) did not estimate the discount factor and did not consider linear adjustment costs. Cooper, Gong, and Yan (2015) estimated an annual discount factor for private plants of 0.929 and found that firing costs exceeded hiring costs.

in the following sections. Here we comment on some key moment and parameter differences.




(a) Data (b) Baseline Model

Figure 1: Employment Growth Response To Innovations: Data and Model
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.

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 the 2000s relative to the 1980s. 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 6.2 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.

The close fit of the model relative to the data is shown through the top row of Figure 1 which isolates the responses to profitability for the two decades in both the data (left panel) and the estimated models (right panel). The reduction in responsiveness in the simulated data matches the patterns in the data closely both qualitatively and quantitatively.

We interpret Table 2 and Figure 1 as providing strong support for our simulated data matching the key properties of the actual data – especially in terms of responsiveness. The tight fit of

Table 4: Decompositions: Intensive and Extensive Margins

| Case | Targeted $\zeta_1$ | $\zeta_2$ |
|---|---|---|
| | 1980s | |
| Baseline 1980s | 0.149 | -0.061 |
| Adj. | 0.135 | -0.041 |
| Prob of Adj. | 0.070 | 0.256 |
| | 2000s | |
| Baseline 2000s | 0.065 | -0.036 |
| Adj. | 0.0445 | -0.0085 |
| Prob of Adj. | 0.058 | 0.220 |

For baseline and adjusters, $(\zeta_1, \zeta_2)$ are the linear and quadratic response of employment growth to the innovation to the profitability shock. For the probability of adjustment, "prob of adj.", we estimate an auxiliary regression using a linear probability model of adjustment for continuers regressed on log(lagged employment) and the profitability shock both linear and quadratic. Reported are the linear and quadratic responses. All moments are from simulated data.

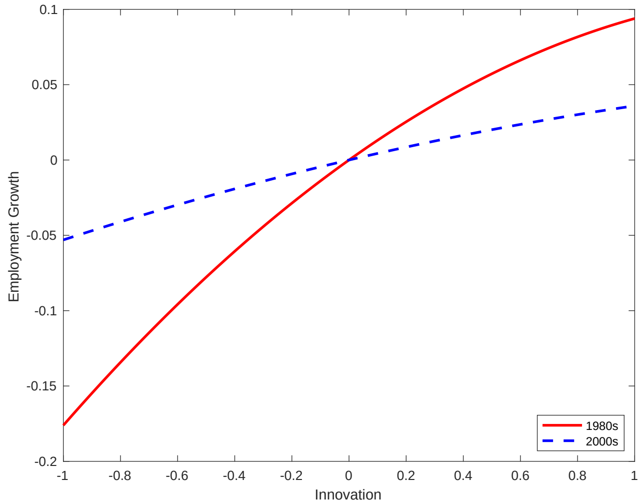
the model to the data gives us confidence to proceed to use the simulated data as a platform for investigating the sources of the decline in dynamism and implications for productivity and markups.[25]

Using the simulated data, we further decompose the growth and innovation relationship for continuing establishments in Figure 1 into the dependence of the probability of adjustment om innovation, the extensive margin, and the relationship between growth and innovation for establishments that adjust, the intensive margin. Table 4 reports this decomposition.

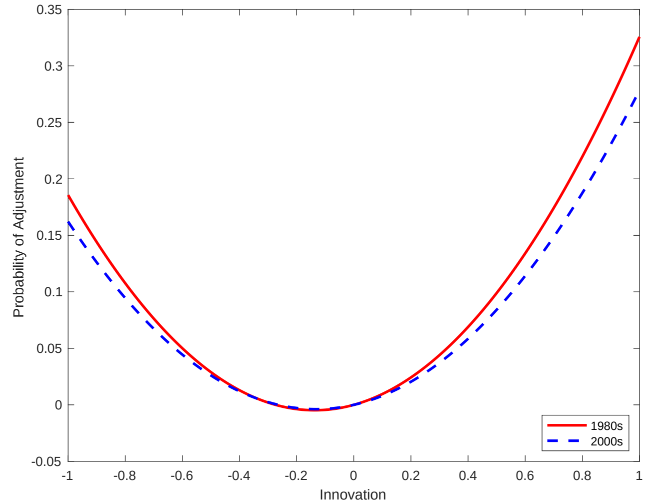
On the intensive margin, it is evident there is a decline in the responsiveness of growth for plants that adjust as both the linear and quadratic coefficients are lower for the 1980s estimates compared to the 2000s estimates. This is seen in the left panel of Figure 2 where the responsiveness in the 2000s is both less sensitive and exhibits less curvature.

On the extensive margin, the regression coefficients are lower in the 2000s simulation results but the differences are not as large as on the intensive margin. From Figure 2, the decline in

[25] Of the nine moments, the one we match the least well is the annual serial correlation of the profitability shocks. Even for this moment the simulated estimates yield a larger decline in serial correlation than is present in the data. This would favor this mechanism as being important as a decline in serial correlation can reduce responsiveness. However, as we show below, the rise in dispersion in both the simulated data (that is matched in the actual data) more than offsets this effect.




(a) Employment Growth Response To Innovations: Adjusters Only




(b) Probability of Adjustment in Response to Innovations

Figure 2: Extensive and Intensive Margins
These figures shows the response to shocks on the intensive and extensive margins.

the likelihood of adjusting is mostly in the tails of the innovation shock distribution. Still, an interesting feature of the probability of adjustment is that it is strongly increasing in the absolute value of the innovation shock.

## 4.3 Relationship Between Parameters and Moments

Before proceeding to the decomposition of the driving forces of declining dynamism exercises, we provide further guidance about identification, with emphasis on the distinct contributions of each of the adjustment cost parameters to matching the moments. There are two exercises: (i) quantifying the response of moments to parameter variations and (ii) selectively setting parameters to zero to study their individuals effects on moments, particularly those representing responsiveness.

Table 5 depicts the elasticities of the moments with respect to parameters evaluated at the 1980s baseline estimates Many of the patterns are straightforward. Increases in all of the adjustment cost parameters decrease the linear term in intensive responsiveness. Increases in the convex component of adjustment costs, $\nu$, increases the quadratic term of the responsiveness while the opposite is true for the nonconvex components.

Key parameters for other mechanisms work in expected ways. An increase in the curvature of the revenue function increases the linear and quadratic terms of intensive margin responsiveness. An increase in the discount factor increases the linear term but decreases the quadratic term. Notable is how sensitive the moments are to small changes in the discount factor. Increases in the persistence and dispersion of shocks increases the linear and quadratic terms of the intensive responsiveness.

From this evidence, all four mechanisms are potential candidates for accounting for the declining responsiveness. However, as is also evident from Table 5, the key parameters also change many other moments. Increases in the convex component of adjustment costs reduces inaction while increases in the nonconvex moments increase inaction. Increases in the curvature of the revenue function, the discount factor and shock dispersion reduce inaction while an increase in persistence increases inaction. The method of simulated moments by construction takes all of these relationships into account.

Table 5: Elasticities

| Parameter | Inact | xrat | $\zeta_1$ | $\zeta_2$ | $\xi_1$ | emp | $\hat{\alpha}$ | $\hat{\rho}$ | $\hat{\sigma}$ |
|---|---|---|---|---|---|---|---|---|---|
| $\beta$ | -9.336 | -33.484 | 390.897 | -223.358 | -26.645 | -17.028 | 9.861 | 17.934 | 5.574 |
| $\nu$ | -0.136 | 0.002 | -1.027 | 1.799 | -0.102 | 0.000 | -0.047 | -0.036 | 0.025 |
| $\gamma_P$ | 0.157 | 0.031 | -0.505 | -2.449 | -0.020 | -0.101 | -0.018 | -0.003 | 0.009 |
| $\gamma_M$ | 0.943 | 0.056 | -1.536 | -3.913 | -0.112 | 0.202 | -0.035 | 0.023 | 0.047 |
| $\Gamma$ | 2.636 | 7.789 | -6.213 | -9.204 | 9.310 | 1.131 | -0.795 | -3.026 | -0.621 |
| $\omega$ | -1.519 | -2.658 | 3.108 | -13.434 | -3.504 | -2.994 | 0.262 | 1.468 | -0.149 |
| $\alpha$ | -0.163 | 0.234 | 1.005 | 3.395 | -0.659 | 13.456 | 0.100 | 1.014 | 0.532 |
| $\rho$ | 0.707 | -19.904 | 6.268 | 7.364 | -1.456 | 0.228 | 0.811 | 3.761 | -0.011 |
| $\sigma$ | -3.429 | -8.317 | 14.603 | 36.604 | -10.355 | -1.407 | 1.234 | 4.136 | 1.825 |

This table shows the response of moments to parameter changes, measured as an elasticity, at the baseline estimates.

As will become clear, it is especially instructive to explore the contribution of each of the adjustment cost parameters to responsiveness. The top part of Table 6 offers an alternative but related perspective showing simulation results starting from the baseline of the 1980s estimate and setting different adjustment costs (indicated in the row) to zero. The bottom part of the table has only those costs.

Table 6: Simulated Moments: Nature of Adjustment Costs

| | Inact | xrat | $\zeta_1$ | $\zeta_2$ | $\xi_1$ | emp | $\hat{\alpha}$ | $\hat{\rho}$ | $\hat{\sigma}$ | £ |
|---|---|---|---|---|---|---|---|---|---|---|
| 1980s Baseline | 0.201 | 0.053 | 0.149 | -0.061 | -0.140 | 10.064 | 0.937 | 0.394 | 0.336 | 1.050 |
| Eliminate adjustment cost component | | | | | | | | | | |
| $\nu$ | 0.261 | 0.043 | 0.571 | 0.134 | -0.136 | 8.555 | 1.085 | 0.457 | 0.334 | 29.603 |
| $\gamma_P$ | 0.094 | 0.040 | 0.403 | -0.353 | -0.121 | 9.037 | 0.996 | 0.431 | 0.334 | 38.217 |
| $\gamma_M$ | 0.054 | 0.027 | 0.706 | -0.575 | -0.103 | 7.952 | 1.020 | 0.470 | 0.327 | 129.949 |
| Retain only this adjustment cost component | | | | | | | | | | |
| $\nu$ | 0.023 | 0.009 | 1.273 | -0.831 | -0.049 | 6.898 | 1.045 | 0.584 | 0.333 | 314.263 |
| $\gamma_P$ | 0.153 | 0.025 | 1.229 | -0.171 | -0.120 | 7.666 | 1.045 | 0.520 | 0.308 | 103.317 |
| $\gamma_M$ | 0.149 | 0.027 | 0.903 | -0.063 | -0.111 | 7.653 | 1.091 | 0.511 | 0.329 | 49.761 |

The moments here are: Inact = $0.025 > \frac{\Delta e}{e} > -0.025$ xrat = exit rate, $(\zeta_1, \zeta_2)$ = linear and quadratic response of employment growth to innovation to profitability shock; $\xi_1$= 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.

From the top panel, eliminating the quadratic adjustment cost increases the curvature of the response, going from concave to convex. Likewise, if the quadratic adjustment cost is the only friction, then the response is convex not concave. From this exercise, it seems that the nonconvex adjustment costs are needed to capture the concavity of the response. However, looking at the bottom panel which treats each adjustment cost in isolation, the relationship is concave if there is only one form of adjustment costs including convex adjustment costs. Also, if there is a single adjustment cost, then the linear part of the response is much higher than the baseline.

A related but distinct inference from Table 6 is that each of the components of the adjustment costs is important as the overall fit worsens considerably relative to the baseline in all cases. The worst fit emerges from keeping only the convex component (bottom panel). The best fit emerges from eliminating only the convex component (top panel). However, this is a relative statement as the fit is much worse than the baseline in this case. Moreover, as noted above, eliminating the convex component makes the intensive responsiveness relationship between growth and innovations convex rather than concave – inconsistent with the data. In contrast, eliminating either nonconvex component makes the intensive responsiveness relationship too concave relative to the baseline. Getting the right magnitudes and curvature of the intensive responsiveness relationship apparently requires both convex and nonconvex components.

These findings relate to the discussion in Ilut, Kehrig, and Schneider (2018) about the source of the concavity of the growth and innovation relationship. Ilut, Kehrig, and Schneider (2018) are agnostic about the source of this observed concavity but indicate it may reflect the pattern of adjustment costs. Table 6 highlights that a mixture of convex and nonconvex adjustment cost components is important for capturing the observed concave relationship.

# 5 Decomposing the Decline in Dynamism

This section decomposes the changes in parameter estimates between the 1980s and the 2000s 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 1980s estimates, allowing other parameters to remain at their estimated values for the 2000s 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 2000s responsiveness moments, holding the others fixed at their estimated values for the 1980s. The model fit is evaluated both in terms of matching the targeted moments as well as the untargeted ones.

### 5.0.1 Simulation Based Decomposition

Table 7 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 the 2000s. 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 1980s values, else parameters are at their 2000s estimated values. So, for example, in the third row of the second block, $\alpha$ is set at its 1980s estimated value of 0.5399 rather than its 2000s estimated value of 0.5231. With all other parameters at their estimated 2000s values, we evaluate the effect on the moments and therefore the fit of the small change in $\alpha$. This exercise is repeated for the other cases.

Table 7: Simulated Moments: Sources of Changes

| | Inact | xrat | $\zeta_1$ | $\zeta_2$ | $\xi_1$ | emp | $\hat{\alpha}$ | $\hat{\rho}$ | $\hat{\sigma}$ | £ |
|---|---|---|---|---|---|---|---|---|---|---|
| 2000s Data | 0.243 | 0.083 | 0.064 | -0.035 | -0.059 | 8.900 | 0.959 | 0.682 | 0.408 | na |
| 1980s Baseline | 0.201 | 0.053 | 0.149 | -0.061 | -0.140 | 10.064 | 0.937 | 0.394 | 0.336 | 1.050 |
| 2000s Baseline | 0.214 | 0.053 | 0.065 | -0.036 | -0.108 | 8.759 | 0.918 | 0.350 | 0.369 | 1.089 |
| $\beta$ | 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 |
| $\alpha$ | 0.219 | 0.056 | 0.052 | -0.014 | -0.109 | 14.630 | 0.914 | 0.356 | 0.373 | 1.894 |
| $(\rho, \sigma)$ | 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 |
| $(\beta, \rho, \sigma)$ | 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 |
| $\nu$ | 0.217 | 0.051 | 0.098 | -0.038 | -0.108 | 8.632 | 0.931 | 0.356 | 0.369 | 1.378 |
| $\gamma_P$ | 0.213 | 0.053 | 0.065 | -0.037 | -0.108 | 8.759 | 0.918 | 0.350 | 0.369 | 1.091 |
| $\gamma_M$ | 0.198 | 0.050 | 0.104 | -0.075 | -0.108 | 8.453 | 0.928 | 0.354 | 0.369 | 2.820 |

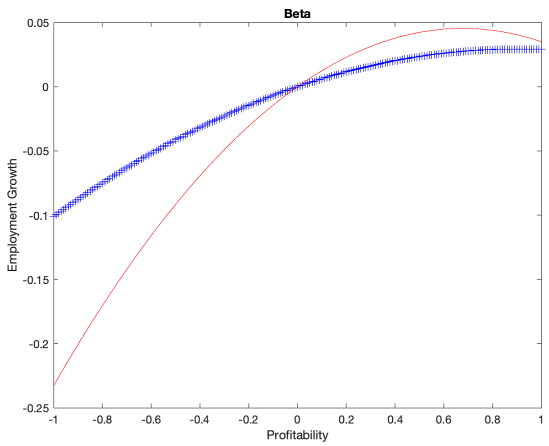
The moments here are: Inact $= 0.025 > \frac{\Delta e}{e} > -0.025$ xrat = exit rate, $(\zeta_1, \zeta_2)$ = linear and quadratic response of employment growth to innovation to profitability shock; $\xi_1$= response of plant-level exit to profitability shock; **emp** is median plant size. $(\hat{\alpha}, \hat{\rho}, \hat{\sigma})$ 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.

The rows of Figure 3 illustrate responsiveness implications for these experiments. The dark curve is the 2000s baseline and consistent with the middle panel of Table 7 the "treatment" comes from holding a particular set of parameters at their 1980s values. If a mechanism is an important part of the change in responsiveness then the *treatment* relationship should move in the direction of the 1980s responsiveness relationship in the right panel of Figure 1.

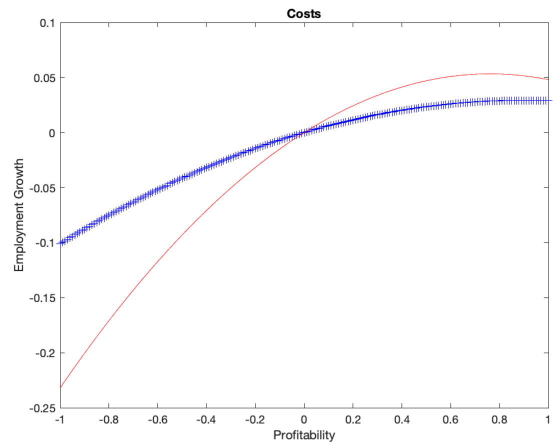
From Table 3, the decline in $\alpha$ goes the right way to induce a decline in responsiveness. However, the estimated decline in $\alpha$ is too small to have much of a quantitative effect. This is evident in the modest change in the fit of the model in Table 7 and the modest effect of this experiment on the responsiveness relationship in Figure 3. In combination, the results don't support changing revenue curvature as a primary source of declining responsiveness.

The experiment for the stochastic processes shows that imposing 1980s values of stochastic processes yields a noticeably worse fit – showing that getting these parameters right is important. However, imposing the 1980s stochastic processes goes the "wrong way" in terms of accounting for changes in responsiveness – the relationship becomes much flatter (and actually declining) rather than steeper. In principle, the estimated decline in $\rho$ contributes to declining responsiveness. However, the decline in $\rho$ from the 1980s to the 2000s is modest and according to Table 5, the increase in $\sigma$ from the 1980s to 2000s works in the opposite direction. Taken together we find that responsiveness is greater in the 2000s at the baseline than that implied by keeping the stochastic processes at the 1980s levels. Table 7 shows that imposing 1980s values of stochastic processes yields an increase in inaction which also goes the wrong way (i.e., inaction in the 1980s was lower than in the 2000s). Given these patterns, we rule out changes in the stochastic process as the source of declining responsiveness.

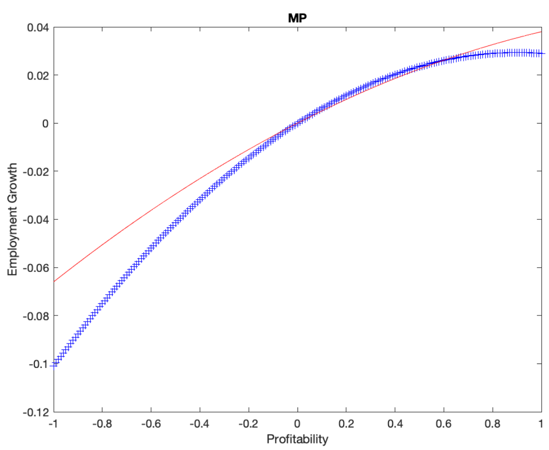
Both the discount factor and the adjustment costs exhibit patterns that suggest they are candidates for accounting for the declining responsiveness. From the upper row of Figure 3, setting the discount factor or the adjustment costs back to their 1980s values increases responsiveness. The discount factor is lower in the 2000s and choices are quite sensitive to this parameter. Keeping $\beta$ at its 1980s value increases the linear part of the responsiveness regression almost back to the 1980s level and increases (in absolute value) the coefficient on the quadratic term. The higher $\beta$ 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



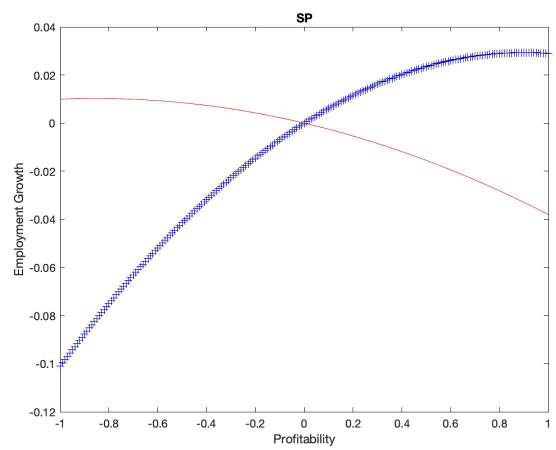

(a) Discount Rate




(b) Adjustment Costs




(c) Market Power




(d) Stochastic Process

Figure 3: Employment Growth Response To Innovations: Decompositions

The figure illustrates the changes in responsiveness from the experiments in Table 7. Dark line is baseline 2000s. Light line sets identified parameter(s) at 1980s values.

the two samples. From Table 7, the deterioration in fit in the treatment comes in large part due to the responsiveness moments. At the 1980s values of the adjustment costs, the two regression coefficients, $(\zeta_1, \zeta_2)$, are much higher than their values in the estimated model for the 2000s and in the data. These results are similar to those from the experiment with $\beta$.

Holding the discount factor at the 1980s estimated value has little impact on the inaction rate while holding the adjustment costs at the 1980s values yields a decline in the inaction rate. On this dimension, the results are more favorable for the adjustment costs mechanism.

The next to bottom panel of the table considers changes in pairs of parameters. Holding both the discount factor and the adjustment costs at their 1980s 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.

To shed further light on the contribution of adjustment costs, the bottom panel of Table 7 repeats this exercise but reverting to 1980s values for each distinct adjustment cost parameter. Reverting to the 1980s value for $\gamma_P$ has little effect while reverting to the lower 1980s value of either $\nu$ or $\gamma_M$ have more pronounced effects – especially $\gamma_M$. The inference we draw is that the increase in both $\nu$ and $\gamma_M$ from the 1980s to the 2000s are the primary sources of the declining responsiveness from increased adjustment costs.

These experiments, summarized by both the moments and Figure 3, suggest that the main contributors to the decreased responsiveness across the decades came from either changes in impatience or in adjustment costs. The analysis in the next subsection provides additional evidence to help distinguish the main source of the reduction in responsiveness

### 5.0.2 Reestimation Based Decomposition

A reestimation based decomposition provides evidence that points to adjustment costs as the primary factor contributing to the reduction in responsiveness – dominating the role of the discount factor. 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 employment 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 8: Targeted Moments: Sources of Changes

| Case | Targeted | | | |
|---|---|---|---|---|
| | $\zeta_1$ | $\zeta_2$ | $\xi_1$ | $\pounds_{targ}$ |
| Data 1980s | 0.113 | -0.054 | -0.081 | na |
| Data 2000s | 0.064 | -0.035 | -0.059 | na |
| Baseline 2000s | 0.065 | -0.036 | -0.108 | 0.691 |
| $\beta$ | 0.097 | -0.035 | -0.164 | 3.433 |
| $C()$ | 0.055 | -0.035 | -0.145 | 2.145 |
| $\alpha$ | 0.140 | -0.088 | -0.144 | 5.779 |
| $(\rho, \sigma)$ | 0.134 | -0.031 | -0.123 | 2.402 |

The moments here are: $(\zeta_1, \zeta_2)$ = linear and quadratic response of employment growth to innovation to profitability shock; $\xi_1$= response of plant-level exit to profitability shock. $\pounds_{targ}$ measures the fit of the model to these moments alone. All moments are from annual data.

For this exercise, Table 8 reports the targeted and data moments while Table 9 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 9. Relative to their 1980s and 2000s 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 1980s baseline. From Table 8 the 2000s 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 9 a value of $\beta = 0.9815$, slightly lower than the original 2000s estimates. This lower value of the discount factor reduces the responsiveness as shown in Table 8. This reduction in $\beta$ creates larger responses to profitability on the intensive and on the exit margins compared to the 2000s data response and so the fit is not nearly as good

[26] For this exercise, we term these targeted moments.

as the adjustment cost case.

For the stochastic process, the serial correlation is about at its 2000s baseline estimate but the re-estimated value of $\sigma$ 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 1980s data.

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

Based on this exercise, the most important contributor to the decline in responsiveness is the rise in adjustment costs. In combination with the simulated based decomposition in the previous section, the rise in adjustment costs is the factor that works well in both sets of decompositions. Reverting to 1980s adjustment costs yields an increase in implied responsiveness, a decline in inaction and a poor fit in the 2000s. Estimating changes in adjustment costs alone leaving other parameters at their 1980s levels and targeting the responsiveness moments yields the best match to the responsiveness moments in the 2000s. No other mechanism works as well in combination across these exercises.

Table 9: Re-estimated Parameters: Sources of Changes

| | $\beta$ | $\nu$ | $\gamma_P$ | $\gamma_M$ | $\Gamma$ | $\omega$ | $\alpha$ | $\rho$ | $\sigma$ |
|---|---|---|---|---|---|---|---|---|---|
| Baseline 1980s | 0.9871 | 4.3716 | 4.7281 | 6.7756 | 0.8652 | 0.1424 | 0.5399 | 0.8785 | 0.5887 |
| Baseline 2000s | 0.9826 | 5.2215 | 4.7473 | 7.4113 | 0.8543 | 0.1257 | 0.5231 | 0.8564 | 0.6288 |
| $\beta$ | 0.9815 | 4.3716 | 4.7281 | 6.7756 | 0.8652 | 0.1424 | 0.5399 | 0.8785 | 0.5887 |
| $C()$ | 0.9871 | 4.1702 | 4.7826 | 9.1077 | 0.8652 | 0.1424 | 0.5399 | 0.8785 | 0.5887 |
| $\alpha$ | 0.9871 | 4.3716 | 4.7281 | 6.7756 | 0.8652 | 0.1424 | 0.5222 | 0.8785 | 0.5887 |
| $(\rho, \sigma)$ | 0.9871 | 4.3716 | 4.7281 | 6.7756 | 0.8652 | 0.1424 | 0.5399 | 0.8565 | 0.6009 |

## 5.1 Interpreting the Increase in Adjustment Costs

The primary message that emerges is that there has been an increase in adjustment costs with both an increase in convex costs and the linear cost of firing. A limitation of this inference is that adjustment costs are themselves a *black-box* and further guidance is needed about the sources of increases in adjustment costs. This sub-section goes beyond the model to discuss possible interpretations of our finding. These interpretations are raised here in the spirit of provoking further study.

**Labor Frictions** Of course, the leading interpretation is that indeed labor frictions have increased. 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. Davis and Haltiwanger (2014) summarize the literature and offer their own evidence that these factors have contributed to declining labor market fluidity including indicators of business dynamism. These analyses are reduced form rather than structural. We leave exploring the connection between these potential sources of increasing adjustment costs and our findings of structural estimates of the changes in adjustment costs 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. For example, our findings suggest that one important component is an increase in firing costs. The changing employment-at-will doctrines in the US is consistent with this pattern. This mapping is interesting and challenging since 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.

In their analysis of four large European countries, Cooper, Horn, and Indraccolo (2023) estimate a dynamic model of labor demand. Looking across countries they find that differences in labor market frictions are a major factor explaining differences in responsiveness to profitability shocks.[27]

An interesting implication of an increase in labor frictions would be the effect on hours per worker. To the extent hour per worker and the number of workers substitute in the production process, an increase in the costs of adjusting workers would imply, *inter alia*, an increased responsiveness of hours to shocks. Decker, Haltiwanger, Jarmin, and Miranda (2020) study this and find that annual hours per worker for production workers did not exhibit much of a change in

[27]Based on a preliminary analysis by Jose Pedro Garcia, policy induced reduction in job protection in Portugal has increased employment responsiveness.

responsiveness. Exploring more comprehensive and higher frequency variation in hours per worker as a margin of adjustment in this context would be an interesting topic for future research.

**Capital Frictions** For tractability, the model excludes capital adjustment costs. Decker, Haltiwanger, Jarmin, and Miranda (2020) present evidence that indeed investment has also become less responsive to innovations. Depending on the interaction of capital and labor, an increase in capital adjustment frictions would lead to an increase in estimated labor frictions in a model that ignores capital frictions. And the logic works in the opposite direction as well: the estimated increase in labor frictions can reduce the responsiveness to investment to shocks as well if capital and labor are sufficiently complementary in the production process.

**Pricing Frictions** The common assumption in sticky price models is that output is demand determined. The pricing decision will impact measures of responsiveness through two channels: (i) the frequency of adjustment and (ii) the nature of the shocks to profitability.

All else the same, if prices do not adjust to a shock, then quantities produced and thus inputs will respond. So clearly, if the degree of price stickiness is lower in the 2000s than in the 1980s, then employment will be less responsive to variations in profitability.[28] Building on Bils and Klenow (2004), sectors are classified based upon their degree of price flexibility. Using the BLS sectoral weights, it is straightforward to calculate how the importance of various sectors evolves over time.[29] Putting these pieces together, it seems that the importance of the sticky price sector has increased by almost 10 percentage points comparing December 1985 with December 2005.[30] This additional rigidity, taken as an exogenous change, would have produced more, not less, employment responsiveness.

Second, without changing the degree of rigidity, a change in the composition of profitability will alter the estimated responsiveness of employment. For these models, if prices do not adjust, then the response of employment will depend on the source of the fluctuation in profitability. Specifically, shocks that increase demand are met by increases in output and hence inputs while

[28] Applying this logic across sectors motivates the research reported in Ozturk and Walsh (2022) which studies the effects of money shocks on investment across sectors using a local projection approach. Importantly, the sectors differ by the degree of volatility in profitability. They find evidence that the real effects of monetary policy are lower in more volatile sectors, where price adjustment is more likely.

[29] See `https://www.bls.gov/cpi/tables/relative-importance/`BLS for a discussion of these weights.

[30] Thanks to Brent Meyer for input on this calculation.

positive productivity shocks lead to (counter) reductions in inputs as demand is fixed given the price. The positive response of employment to profitability could be seen as evidence that demand shocks prevail. Still, technology shocks could play a larger role in the 2000s, either because these shocks were a larger part of the total variation or became more persistent. This would lower the overall responsiveness of employment to profitability given that firms, conditional on non-adjustment, would respond in an inverse way to technology shocks. We have no evidence of this at the monthly frequency corresponding to that of price measurement.

# 6 Implications for Productivity and Markups

This section uses the estimated model to look at productivity and markups over the two samples. One point is to evaluate the productivity implications of the estimated increase in adjustment costs. The second is to explore the effects of the increase in adjustment costs on measured markups.

## 6.1 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.[31]

Table 10 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 (the sum of revenue across all firms divided by the sum of employment across all firms), a measure of aggregate productivity, (AggProd), (ii) the time series mean of the cross-sectional (across firms) standard deviation of the average revenue product of labor (Mstd), (iii) the time series mean of the cross-sectional (across firms) covariance between the profitability shock and employment, $c(A, e)$.

Table 10 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

[31] Interestingly, 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.

Table 10: Productivity Implications

| Sample | AggProd | Mstd | $corr(A,e)$ |
|---|---|---|---|
| 1980s | 1 | 7.147 | 0.768 |
| 2000s | 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 1980s as a normalization.

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 our measure of aggregate productivity.[32]

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.[33] Our results suggest 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.

## 6.2 Markups

The distribution of markups in our model is, by the specification of constant elasticity demand functions, degenerate. Further, 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. These findings contrast with those in relative studies. Prominently, 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.

[32] Our analysis abstracts from general equilibrium effects influencing aggregate output and input prices, particularly wages. We impose some discipline on the real wage by targeting median size and the exit rate. Our focus is on the implications of increased within-industry misallocation induced by rising adjustment costs. All three of the simulated moments we target reflect this increased misallocation.

[33] For 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 1980s and the 2000s. Our model environment only has one factor so this is equivalent to total factor productivity in our model setting. We have not considered the possible increase in capital adjustment costs over this period.

Some aspects of our results are not necessarily inconsistent with the De Loecker, Eeckhout, and Unger (2020) findings for a couple of reasons. 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, De Loecker, Eeckhout, and Unger (2020) find that much of the increase in the revenue-weighted markup reflects reallocation towards large revenue firms that have higher markups. Since our analysis has no dispersion in actual markups across firms, at first glance it appears we are not capturing this important feature of their findings and potentially important part of the impact of changing markups on responsiveness. However, while our model yields no dispersion in actual markups, our model yields dispersion in the measured markups using the De Loecker, Eeckhout, and Unger (2020) measurement approach.

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 $\mu_{it}$ is the measured markup, $\theta_{it}$ is the output elasticity of labor and $ls_{it}$ is the share of **total revenue** that is paid to labor.[34] 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 within subperiods, 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 $\theta_{i,1980s} = 0.673$ 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 1980s and 2000s decades

[34] Bond, 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.

(including the estimated variation in $\alpha$).

Results are reported in Table 11. 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. They highlight that the increase in the gap between the $90th$ and median is an indicator of rising dispersion and skewness in the revenue-weighted distribution of markups.

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 the 1980s and 2000s. This is in spite of there being no cross-sectional variation in actual markups in our framework and only a modest reduction in $\alpha$ across the two 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.[35] In terms of skewness, De Loecker, Eeckhout, and Unger (2020) don't report the 10th percentile but this is easy to calculate in our simulated data. Using this calculation we can compute a measure of changing skewness given by the change in the difference between the $90 - 50th$ percentile and the $50th - 10th$ percentile. We find that this difference increases from 0.07 to 0.15 from the 1980s to 2000s implying an increase in not only dispersion but skewness.

The bottom panel of Table 11 provides some counterfactuals intended to decompose the changes in the markup distribution over time by setting some parameters back to their 1980s values. From these results, there is no single mechanism that explains the findings. Holding adjustment costs

[35] Foster, 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.

at their 1980s estimates reduces the mean markup slightly. The lower estimate of $\alpha$ in the 2000s plays a slightly larger role: setting this curvature at its 1980s 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 1980s value almost reproduces the 1980s findings.

Table 11: Moments of Measured Markups

| | Mean $\mu$ | Median $\mu$ | $P90$ $\mu$ | $Corr(\mu, \frac{R}{\sum R})$ | $Corr(\mu, A)$ |
|---|---|---|---|---|---|
| | | | 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 1980s est. | | | 2000s Decompositions | | |
| $C(\cdot)$ | 1.67 | 1.61 | 2.39 | 0.23 | 0.50 |
| $\alpha$ | 1.64 | 1.57 | 2.36 | 0.17 | 0.47 |
| $\beta$ | 1.67 | 1.61 | 2.42 | 0.27 | 0.54 |
| $\rho, \sigma$ | 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)).[36] 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.

[36] De Loecker, Eeckhout, and Mongey (2021) generate heterogeneous markups in their model using a model with oligopolistic competition that features higher markups for larger, more productive firms.

# 7 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, they were less able to do so while simultaneously matching other moments. The increase in adjustments costs itself seems to be broad based, including both quadratic as well as the linear costs of firing workers.

Increased costs of labor adjustment yield aggregate 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 1980s 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 measured markup distributions. 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 employment 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, 135, 645–709.

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.

Bils, M., and P. J. Klenow (2004): "Some evidence on the importance of sticky prices," Journal of political economy, 112(5), 947–985.

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.

Davis, S. J., J. C. Haltiwanger, and S. Schuh (1998): "Job creation and destruction," MIT Press Books.

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, 110, 3952–3990.

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: Implicatios 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.

——— (2021): “The micro-level anatomy of the aggregate labor share decline,” Quarterly Journal of Economics, 136(2), 1031—-1087.

Ozturk, O., and T. Walsh (2022): “Sectoral Volatility and the Investment Channel of Monetary Policy,” Available at SSRN 4009725.

Wooldridge, J. M. (2009): “On Estimating Firm-Level Productivity Functions Using Proxy Variables to Control for Unobservables,” Economics Letters, 104(3), 112–114.

# Data Appendix

Our core moments build on the data infrastructure from Foster, Grim, and Haltiwanger (2016) and Decker, Haltiwanger, Jarmin, and Miranda (2018) (an an earlier working paper version of Decker, Haltiwanger, Jarmin, and Miranda (2020)) integrating the ASM and the LBD from from 1980 to 2010. We use the ASM to generate measures of profitability shocks and we use the LBD to construct measures of employment growth for continuing and exiting manufacturing establishments. Importantly, 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. We follow the measurement and timing conventions of Decker, Haltiwanger, Jarmin, and Miranda (2018). In this respect, the representative sample with propensity score weights for the analysis of responsiveness with respect to innovations follows the approach in that paper (and the other cited papers in this discussion). As noted in the text, the reponsivness regressions using the actual data permit the coefficients to vary by year using a quadratic trend. We use the estimated coefficients from those specifications to compute responsiveness effects by year and then take decade averages of those annual responsiveness effects.

The LBD is used for the inaction and median moments. The exit rate moments are based on the Business Dynamic Statistics (BDS) that are in turn based on the LBD. The moments in Table 2 are decade averages of annual statistics. The annual statistics have been adjusted for time varying cyclicality in a manner similar to the cyclical adjustments in our responsiveness regression (see the discussion of controls in that specification). This adjustment does not matter much in practice but the objective is for the differences in decade averages to reflect low frequency variation that is the focus of this paper.