Online Appendix: The Micro and Macro of Managerial Beliefs

A Data Appendix

A.1 Representativeness of the SBU

Figures A.7 to A.10 explore how representative SBU data are of the broader US Economy. Similar figures appear in Altig et al. (2018), each showing the share of employment accounted for by firms in different size categories, ages, sectors, or regions in the Survey of Business Uncertainty and the US Economy. To compute employment shares by size, sector, and region I use the US Census Bureau’s Statistics on US Businesses for 2015. For firm age I use the Census’ Business Dynamics statistics.

The SBU is more or less representative of the US in employment-weighted terms. In Figure A.7 we can see the share of employment accounted for by firms with more than 500 employees is somewhat higher for the SBU, as is the share of employment born prior to 1990 in Figure A.8. The survey’s industrial composition includes firms from all sectors, as we can see in Figure A.9, but relatively more firms in durables manufacturing and finance and insurance, and relatively fewer in the health care sector. Deviations in the share of employment for particular groups of firms versus the US economy stem partly from the survey’s Dunn & Bradstreet sampling frame, in which smaller and younger firms may be under-represented. It may also be that older and larger firms are more willing to respond to the survey.

A.2 Measuring Subjective Moments and Forecast Errors in the SBU

My analysis focuses on managers’ subjective beliefs about the growth rate of quarterly sales looking four quarters into the future as reported in the SBU. Managers and analysts pay significant attention to sales and earnings as indicators of firm-level performance, so focusing on sales growth as a performance outcome seems a reasonable choice. Anecdotally, managers also use sales figures and projections for planning, budgeting, and hiring. As I report in Appendix A.9, my three facts concerning managers’ beliefs about sales growth are also present in their beliefs about employment growth looking 12 months ahead, suggesting they are a robust feature managerial beliefs rather than a one-off finding.

The SBU is an unbalanced monthly panel, and an individual respondent receives the questionnaire that includes questions about sales every two months\(^{34}\). Throughout my analysis of managers’ beliefs in Section 2 I preserve the survey’s structure as a monthly panel with gaps, but the results are similar if I collapse the panel to a quarterly frequency, picking the last response of the quarter. Indeed, for my structural estimation exercise in Section 4.2, I compute the target moments that come from the SBU using the last response of each calendar quarter for conformity with my quarterly model.

A.2.1 Measuring subjective moments

I focus on SBU respondents’ subjective beliefs for sales growth over the next four quarters. Responses from prior to September 2016 report subjective probability distributions over the dollar level of sales looking four quarters ahead. Thus, I first compute the rate of sales growth implied by the respondent’s reported sales level in quarter \(t\), \(s_t\), and each of the five potential quarterly sales levels \(s_{j,t+4}\) for \(j = 1, 2, 3, 4, 5\).\(^{35}\) Following the convention in the literature on business dynamics, I measure these five potential growth rates as the difference across periods divided by the average:

\[
g_{j,t+4} = \frac{s_{j,t+4} - s_t}{\frac{1}{2}(s_{j,t+4} + s_t)}. \tag{9}
\]

\(^{34}\)Since the last major survey update in September 2016, respondents receive questionnaires about sales and employment in one month and then questionnaires about capital expenditures and unit costs the next month. Prior to September 2016, the SBU also asked questions about pricing and profit margins, so respondents received the same questionnaire approximately once every three months.

\(^{35}\)For simplicity of notation I do not use respondent-level subscripts throughout this section, but responses in the SBU belong to a respondent manager \(i\) in month \(m\) which belongs to quarter \(t\)
Survey responses from September 2016 and later report subjective distributions over sales growth rates directly. I assume a respondent’s estimate for her firm’s sales growth rate between quarters \( t \) and \( t+4 \) under scenario \( j = 1, 2, 3, 4, 5 \), \( x_{j,t+4} \), corresponds to a traditional growth rate, so that \( x_{j,t+4} = (s_{j,t+4} - s_t)/s_t \). Here again, \( s_t \) is the current sales level and \( s_{j,t+4} \) is the potential sales level in quarter \( t+4 \) under scenario \( j = 1, 2, 3, 4, 5 \). Therefore, I translate these raw data growth rates to conform with the formula in 9, measuring

\[
\begin{align*}
g_{j,t+4} &= \frac{2 * x_{j,t+4}}{2 + x_{j,t+4}}.
\end{align*}
\]

For each survey response, I now have a five-point subjective probability distribution over sales growth between quarters \( t \) and \( t+4 \). These subjective distributions consist of a potential growth rate for each of 5 scenarios \( \{g_{j,t+4}\}_{j=1}^{5} \) and their associated subjective probabilities, \( \{p_{j,t+4}\}_{j=1}^{5} \). I then calculate a firm’s forecast for her firm’s sales growth between quarters \( t \) and \( t+4 \), \( g_{t+4} \) is:

\[
\tilde{E}_t[g_{t+4}] \equiv \tilde{E}_t[g_{t+4}|I_t] = \sum_{j=1}^{5} p_{j,t+4}g_{j,t+4}.
\]

Here \( \tilde{E}_t[\cdot] \) is my notation for the respondent’s subjective expectation on date \( t \) when she responded to the survey. I use \( I_t \) to denote the managers’ information set at \( t \), which includes other information reported in the survey – for example the firm’s current sales level \( s_t \) – and other information about the firm’s current and future prospects that may be known to the manager but is unobservable to researchers. Throughout my analysis I assume managers’ subjective beliefs, and thus any moments based on those subjective distributions condition on \( I_t \), as do any ex-post realized outcomes I observe.

I compute other moments of managers’ subjective distributions by analogously taking the inner product of managers’ subjective probabilities and some function of potential sales growth outcomes. For example a manager’s subjective mean absolute deviation from her forecast (on average how far she expects to be from her forecast) is:

\[
\tilde{\text{MAD}}_t[g_{t+4}] = \tilde{E}_t[\sum_{i=1}^{5} \tilde{p}_{j,t+4} \left| g_{j,t+4} - \tilde{E}_t[g_{t+4}] \right|] = \sum_{i=1}^{5} \tilde{p}_{j,t+4} \left| g_{j,t+4} - \sum_{k=1}^{5} \tilde{p}_{k,t+4}g_{k,t+4} \right|. 
\]

Parts of Section 2 use managers’ subjective mean absolute deviations as a measure of their expected forecast accuracy, or equivalently their ex-ante uncertainty.

### A.2.2 Realized sales growth and forecast errors

I measure a respondent’s firm’s actual sales growth between quarters \( t \) and \( t+4 \), \( g_{t+4} \) based on the respondent’s reported sales level in quarter \( t \) when she makes her forecast as well as four quarters later. Specifically, I exploit the fact that the \( SBU \) is a panel, and measure this ex-post realized sales growth as

\[
g_{t+4} = \frac{s_{R,t+4} - s_t}{\frac{1}{2}(s_{R,t+4} + s_t)}.
\]

\(36\) Here I am making some abuse of notation. Although \( \tilde{E}_t[\cdot] \) does denote the respondent’s subjective expectation at the time of the survey, I may observe multiple forecasts for sales growth between quarters \( t \) and \( t+4 \) if the respondent answered questions about sales in more than one month belonging to quarter \( t \). Again, restricting my analysis to the final response of the quarter does not materially change any of the main results.
where $s_t$ is the firm’s sales level as reported in quarter $t$, and $s_{R,t+4}$ is the realized sales level in quarter $t+4$. Although seemingly straightforward, the fact that the SBU asks questions about sales every two months (three months prior to September 2016) means that there may be more than one forecast and more than one reported sales level within a calendar quarter. Individual respondents may also drop out of the sample or fail to respond to the survey in a particular month.

To accommodate these circumstances, I aim to measure sales growth realizations based on sales levels reported exactly twelve months apart. If I do not observe a sales level response exactly twelve months after observing the original sales level $s_t$, I proceed as follows:

- If $s_t$ belongs to the first month of the quarter (e.g. January), I record $s_{R,t+4}$ based on the sales level thirteen months after observing $s_t$. If there is also no sales level reported thirteen months later I use the sales level fourteen months after.
- If $s_t$ belongs to the second month of the quarter (e.g. February), I record $s_{R,t+4}$ based on the sales level eleven months after observing $s_t$. If there is no sales level reported eleven months later I use the sales level thirteen months after.
- If $s_t$ belongs to the third month of the quarter (e.g. March), I record $s_{R,t+4}$ based on the sales level eleven months after observing $s_t$. If there is no sales level reported eleven months later I use the sales level ten months after.

Following this procedure, I increase the number of sales realizations I observe four quarters after recording a subjective distribution, namely because I don’t require exactly twelve months between beliefs and realizations. But I am also careful to record the realized level of quarterly sales in the appropriate quarter, that is quarter $t+4$ for a subjective distribution recorded in quarter $t$.

Having obtained data on realized sales levels $s_{R,t+4}$, I can compute the realized growth rates $g_{t+4}$ and managers’ forecast errors by taking the difference between their sales growth forecast (i.e. their subjective mean sales growth for quarters $t$ to $t+4$) and the ex-post sales growth realization:

$$ForecastError_{t,t+4} = \tilde{E}_t[g_{t+4}] - g_{t+4}.$$  

Using this definition, a positive forecast error occurs when a respondent’s subjective mean exceeds the realized sales growth over the ensuing four quarters, and vice versa for a negative forecast error. For much of my analysis in 2 I winsorize forecast errors at the 1st and 99th percentiles to limit the influence of outliers but my results are similar without winsorizing.

### A.3 Sample selection and descriptives

My main sample in 2 consists of all firm-months for which I can construct a forecast error for sales growth looking four quarters ahead. Namely I include in my sample all responses to the sales module in the SBU with well-formed probability distributions for which I can observe a subsequent sales level realization four quarters later.

Following Altig et al. (2018), I review any forecast errors with magnitude greater than unity\(^{37}\) and correct units mistakes (e.g. reporting $5 instead of $5,000,000 in quarterly sales) or other common patterns (e.g. reporting annual rather quarterly sales) based on the history of responses. After this review, I exclude forecast errors for which there is no obvious mistake and are significantly larger than one in absolute value. To conduct this review I exploit the fact that respondents provide estimates of their firm’s current quarterly sales every two months, so looking at this history of responses I can easily determine whether individual responses appear anomalous in comparison with the months immediately before and after.

\(^{37}\)Note that under my sales growth measure $g_{t+4}$ is bounded by plus and minus two. A forecast error equal to positive one will arise, for example, if the firm predicts sales growth of zero and the firm’s sales subsequently drop by two-thirds.
The sample I use in the rest of this empirical section is an unbalanced monthly panel with gaps, including 1,574 forecast error observations pooling across all firms and months. Prior to September 2016 (when firms answered questions about sales only every three months), my sample contains about 20 to 30 forecast errors per month, while months since September 2016 have about 100 forecast errors per month.

Again, the firms who have forecast errors are larger, well-established organizations, covering a mix of privately-held and publicly-traded firms. Table A.1 displays basic summary statistics about the sample. The median and mean employment in the sample are 175 and 477, the median and mean quarterly sales are $8 Million and $38.1 Million. See Figures A.3 and A.4 for histograms of the distribution of current employment and current quarterly sales across all forecast error observations. The firms in the sample are also fairly old, with about half of all forecast errors belonging to firms that hired their first paid employee prior to 1970, and only 3 percent since 2000, as we can see in Figure A.5.

Performance across the sample is highly heterogeneous, as we should expect from a large cross section of firms. While the mean sales growth over four quarters is 0.046 (on the DHS scale ranging from -2 to +2), its standard deviation is 0.270. Figure A.6 shows the distribution of four-quarter sales growth realizations for my main sample.

A.4 How do Beliefs About Sales Growth in the SBU Relate to Outcomes?

This section elaborates further on my results from Section 2.2 about whether managerial beliefs about future outcomes from the SBU can predict future outcomes and decisions. These exercises aim to rule out the hypothesis that the beliefs recorded in the survey are mostly or all noise. If that were the case, my empirical results arguing that managers have biased beliefs would carry little weight. It would be hard to argue that my data show managers are biased if the data is too noisy to represent beliefs in the first place.

I additionally want to rule out the case that managerial forecasts for sales growth are completely unrelated to current hiring and hiring plans. My quantitative analysis assumes that managers employ the beliefs they provide in the survey when they make business decisions, which is why biases affect decisions. If current hiring and hiring plans are orthogonal to sales growth forecasts and uncertainty, it is hard to argue that the biases I identify in the SBU affect businesses’ decisions.

A.4.1 Realized sales growth and employment vs. their ex-ante forecasts

Figure A.12 reproduce the results shown in the main text, namely that sales and employment growth forecasts (where a forecast equals the mean of a manager’s subjective distribution) have extremely high predictive power for actual sales and employment growth over the next four quarters or twelve months, as appropriate. The t-statistics for these regressions, based on firm-clustered standard errors, are respectively 4.9 and 7.7, confirming the visual intuition from the figure that ex-ante subjective forecasts can predict realized outcomes. Although in both cases the coefficient in the implied regressions is less than one, this may be attenuation bias due to classical measurement error in the forecast. Previous literature typically finds highly significant coefficients (often less than one) in these sorts of regressions, for example in Gennaioli et al. (2016). My findings are also consistent with an older literature studying the subjective beliefs of households, which shows that people are willing and able to answer surveys and that their answers are highly predictive of future events (see Manski, 2004 and 2018, for a survey of this literature).

A.4.2 Sales growth forecasts predict planned hiring

Figure A.13 shows that hiring plans (i.e. managers’ forecast for employment growth over the next twelve months) correlate highly with with their sales growth forecasts for a four-quarter horizon, with a t-statistic of 8.5. This result is highly encouraging, as it means managers’ sales forecasts are highly consistent with their best guess for their own actions looking ahead to the next year. Given we already saw in Figure A.12 that these hiring plans are also highly predictive of actual hiring, this is even more evidence that managers’ beliefs as stated in the SBU are an important input into hiring plans and hiring decisions.

---

38The SBU asks managers for beliefs about sales growth over a horizon of four quarters, while the horizon for employment questions is twelve months.
A.4.3 Actual sales growth predicts actual hiring

Figure A.14 validates that actual hiring and actual sales growth are correlated in the SBU. Namely, when firms record positive sales growth they also hire more workers as we would expect in most economic models. This is an additional validity check for the data, having shown that sales and employment growth forecasts are consistent and predict realizations.

A.4.4 Actual hiring versus forecast sales growth

In Figure A.15 I show that there is a positive but potentially nonlinear relationship between actual hiring in the year after managers’ make a sales growth forecast and the forecast itself. The weak linear relationship may be due to nonlinearities in the relationship between sales growth forecasts and actual hiring decisions that may take place throughout the year following the forecast. Given my earlier findings that forecasts predict hiring plans and hiring plans predict actual hiring, and finally sales growth forecasts predict actual sales growth I view this as consistent with the need to use a nonlinear model to understand how sales growth forecasts ultimately result in hiring decisions.

A.4.5 Current hiring versus sales growth forecasts and current sales growth

Figure A.16 explore how hiring in the current quarter relates to managers’ beliefs about sales growth looking ahead over the next four quarters, and the firm’s recent performance. In Figure A.16a we can see that there is a weak but positive relationship between the firm’s current net hiring and managers’ medium-run sales growth expectations. Turning to Figure A.16b, there is by contrast a clear relationship between hiring in quarter \( t \) and the firm’s performance in quarter \( t \) relative to quarter \( t - 1 \). This pattern suggests managers’ current decision to hire or lay-off workers incorporates several different pieces of information, including both their medium-run forecasts for the firm’s performance and also the firm’s current business conditions.

That innovations in the firm’s sales predict their hiring more strongly than longer run forecasts resembles the dynamics in my estimated model in the main text, in which overconfident and overextrapolative managers are constantly reacting to changes in their firms’ business conditions. Furthermore, I interpret the weaker relationship between expectations and hiring this quarter relative to hiring plans for the next twelve months as evidence that hiring is not frictionless, but rather subject to adjustment costs and frictions that are at the center of the economic model I use to quantify the impact of beliefs biases.

Based on the evidence in this subsection, I therefore argue managers’ beliefs about their firm’s future performance do enter their current hiring decisions and – more importantly – their hiring plans. These results are consistent with the evidence in the literature on beliefs and expectations, in particular the finding by Gennaioli et al. (2016) that survey-based expectations can predict outcomes and actions like investment above and beyond model-implied proxies like Tobin’s \( q \). However, since my SBU data is confidential and mostly comes from privately-held firms, so I cannot at the moment link it to rich data sources to test whether beliefs predict outcomes above and beyond other variables.

A.5 Managerial Overconfidence Underestimates the Level of Risk, Rather than Differences in Risk

How does managerial overconfidence manifest itself for firms that experience more versus less ex-ante subjective uncertainty? I find managers underestimate an approximately fixed level of risk regardless of their firm’s subjective uncertainty, while being highly sensitive to differences in risk across firms. To see this, I explore the relationship between ex-ante uncertainty and ex-post absolute forecast errors in Figure A.18. The horizontal axis shows 20 equally sized quantiles of subjective uncertainty – the standard deviation of managers’ subjective distribution for sales growth between \( t \) and \( t + 4 \). The blue circles on vertical axis show the average absolute forecast error that arises empirically in each of those twenty quantiles, while the orange triangles show what those average forecast errors would look like if sales growth realizations were distributed according to managers’ subjective distributions.

The two bin-scatters in Figure A.18 are upward sloping and essentially parallel, showing that managers
who report higher ex-ante subjective uncertainty make larger ex-post errors both empirically and under the subjective distribution (the latter by construction). This result is consistent with existing work, including Bloom et al. (2017), who find that managers in industries and firms with higher historical and option-implied volatility report higher uncertainty. The similarity in the slope of the relationship across the empirical and subjective errors implies that differences in risk as perceived across managers reflects differences in true risk fairly accurately. But the vertical gap between the empirical and subjective errors in Figure A.18 visually captures the degree of overconfidence, which appears to be constant. At all levels of uncertainty, managers appear to underestimate the magnitude of their forecast errors by a fixed amount. To my knowledge, this is the first paper to document that managers seem to underestimate a fixed level of risk, while remaining sensitive to differences in risk across firms and time.

Repeating this exercise focusing on within-firm variation in subjective uncertainty the results are similar, but the slope of the relationship between empirical absolute forecast errors and subjective uncertainty decreases by about half. This decrease in the slope suggests that managers are better at recognizing differences in subjective uncertainty across rather than within firms.

A.6 Overconfidence or Measurement Error?

In this section I argue that measurement error is unlikely to be responsible for the large excess absolute forecast errors I find in the SBU data. Recall that the mean excess absolute forecast error is the key statistic I use to quantify the degree of managerial overconfidence. My measures of realized sales growth in the SBU are almost certainly measured with error, for example because managers report their firm’s current sales level in the SBU in round numbers before official accounting figures are published. It’s true that SBU respondents don’t have strong incentives to be perfectly accurate, and may answer the survey without paying much attention. The question is whether these potential sources of measurement error can explain why I measure large forecast errors (in absolute terms), while managers’ subjective distributions imply they should be making much smaller errors.

To test whether the magnitude of my measured forecast errors seems implausible, in Figure A.19a I compare the distribution of forecast errors for sales growth I obtain from the SBU, against the distribution of errors made by professional analysts in IBES, both from a horizon of four quarters. We should expect sales data released by public firms to have significantly less measurement error than there might be in the SBU, given realized values come from official accounting releases that are meant for public distribution. They definitely should not suffer from rounding, units, and inattention issues that we might worry about in the SBU data.

Owing to the structure and variables available in IBES I construct forecast errors somewhat differently from the main analysis in the paper. IBES reports forecasts and realizations of the level of sales. Therefore I construct implied forecasts for the growth rate of sales (looking four quarters ahead) by taking the growth rate implied by the current level and the forecast. In the SBU I compute the implied subjective expectation for the sales level four quarters ahead and then the growth rate implied by that expected future level with the current reported level. Then I define forecast errors in both the SBU and IBES as the difference between these growth forecasts and actual growth.

Looking at Figure A.19a, it seems that managers do make slightly larger forecast errors in the SBU than do analysts in IBES. This may be partly due to additional measurement error in the SBU, but it is also a well-known fact (see, for example Davis et al., 2007) that larger firms are less volatile. Firms in IBES are some of the larger publicly-traded firms in the US so makes sense that their sales to be more predictable than the sales of firms in my SBU sample of smaller (though still fairly large in absolute terms) firms. In Figure A.19b I confirm that the distribution of SBU forecast errors under managers’ subjective distributions for realized sales growth look implausible. There I show the distribution of empirical forecast errors from IBES and the subjective distribution of forecast errors in the SBU. As in Figure 5 in the main text, it is the subjective distribution of errors that looks implausible in comparison with the distribution of actual analyst errors in IBES.
A.7 Is Overconfidence Mechanically Generated from the SBU’s Five-Point Subjective Probability Distributions?

Responses in the SBU elicit managers’ beliefs about future sales growth using a five-point discrete distribution. Realized sales growth, however, is a continuous variable. This discrepancy raises the concern that my key measures of overconfidence − excess absolute forecast errors − may be mechanically large because discrete approximations simply have a hard time capturing continuous distributions. I argue that is not correct.

Looking at Figure 5 in the main text, it is clear that the managers’ subjective distributions overestimate the probability of small forecast errors, not that actual forecast errors are occasionally very large because the distribution of sales growth realizations has continuous support. Managers place nearly 75 percent probability on the possibility that forecast sales growth will be within 5 percentage points of realizations. Empirically this only happens with about 25 percent probability. Managers in turn underestimate the probability of being off by about 10 percentage points, which is actually very much within the realm of normal (the standard deviation of actual forecast errors is close to 0.25). These patterns suggest that managers place the five bins corresponding to the lowest, low, middle, high, and highest scenarios too close together, ignoring a large amount of mass at the tails of the true distribution.

To verify this intuition I demonstrate that making a discrete five-point approximation of the continuous distribution of realized sales growth, ignoring modest amounts of mass at the tails, need not generate large excess absolute forecast errors. I consider two potential approaches. Under a first approach based on the Tauchen (1986) algorithm, I first pick some amount of tail mass \( p \in (0, 1) \) of the empirical distribution to disregard. Then I pick five equidistant bins \( q_i \), \( i = 1, 2, 3, 4, 5 \) where \( q_1 \) and \( q_5 \) are the endpoints of remaining support. Finally, I distribute probabilities across bins based on the cumulative distribution \( F(\cdot) \) of realized sales growth:

\[
\begin{align*}
p_1 &= F \left( \frac{q_1 + q_2}{2} \right) \\
p_2 &= F \left( \frac{q_2 + q_3}{2} \right) - F \left( \frac{q_1 + q_2}{2} \right) \\
&\vdots \\
p_5 &= 1 - F \left( \frac{q_4 + q_5}{2} \right). 
\end{align*}
\]

Once I have this approximate discrete distribution I use it to construct a forecast \( E = \sum_{i=1}^{5} p_i q_i \) and a mean absolute deviation \( \text{MAD} = \sum_{i=1}^{5} p_i ||q_i - E|| \). Then I find the mean absolute forecast error implied by the discretization, \( \text{MAFE} = \sum_{n=1}^{N} (E - g_n) \) where \( n \) index observations of the empirical sales growth distribution I am targeting. Finally I find the excess absolute forecast error generated by my discrete approximation, \( EAFE = \text{MAFE} - \text{MAD} \), which is analogous to my measure of overconfidence in Section 2. Table A.2a shows how this excess error changes across discretizations that ignore the outermost \( p \) mass of the target empirical distribution. Ignoring modest amounts of tail mass \( (p \leq 0.2) \) results in modest excess absolute forecast errors, on the order of a couple of percentage points. Ignoring the outermost 40 percent of the mass (i.e. placing the outermost bins at the 20th and 80th percentiles of the target distribution) only generates an excess absolute forecast error half as large as the excess error I observe empirically.

I find similar results using an alternative approach to discretizing the empirical distribution. The mean probability vector SBU respondents assign is approximately \( (p_1, p_2, p_3, p_4, p_5)' = (0.1, 0.2, 0.4, 0.2, 0.1)' \), an intuitive, unimodal, and symmetric distribution. This second method uses these probabilities and the corre-

\[39\] Indeed, the literature that works with discrete-time dynamic programming models has used discrete approximations to Gaussian Markov processes at least since Tauchen (1986) without major concerns that the discrete approximations mechanically understate the dispersion of the stochastic process in question.

\[40\] To be specific, my target distribution is the empirical distribution of sales growth realizations purged of heterogeneity due to differences in managers’ subjective expectations. Purging this heterogeneity involves regressing realized sales growth on SBU managers’ ex-ante forecasts and working with the residual from this regression. Residualizing ensures that the variance of my target continuous distribution reflects unpredictable variation in realized sales growth rather than predictable variation that managers have in their own information sets.
sponding quantiles of the empirical distribution (again, disregarding some of the outermost mass \( p \in (0, 1) \)) to select the five support points of the discrete distribution. Namely, pick \( q_i, \ i = 1, 2, 3, 4, 5 \) such that:

\[
\begin{align*}
p_1 &= F\left(\frac{q_1 + q_2}{2}\right) \\
p_2 &= F\left(\frac{q_2 + q_3}{2}\right) - F\left(\frac{q_1 + q_2}{2}\right) \\
&\vdots \\
p_5 &= 1 - F\left(\frac{q_4 + q_5}{2}\right)
\end{align*}
\]

where \( F(\cdot) \) now is the CDF of the target distribution ignoring the pre-determined \( p \) mass. I can then use this discretization to construct measures of excess error \( EAFE \) as for the "Tauchen" approach. Table A.2b shows that using this quantile-based approach also does not mechanically generate large excess absolute forecast errors when ignoring modest amounts of tail mass \( p \). For \( p = 0.4 \), the implied excess error is even smaller at 0.058 than for the "Tauchen" approach, effectively ruling out the hypothesis that discretization on its own can generate the large excess errors we find in the SBU data.

### A.8 Additional Evidence of Overextrapolation

In this section I show some additional evidence consistent with my claim that managers responding to the SBU overextrapolate from current conditions when they form beliefs about future sales growth.

#### A.8.1 Breaking down the relationship between forecasts, realizations, and recent sales growth

I confirm that managers appear to underestimate mean reversion of short-term shocks by looking at Figure A.20, in which I repeat the bin-scatter from Figure 7, now plotting forecast and realized sales growth separately on the vertical axis. Managers’ forecasts for sales growth between quarters \( t \) and \( t + 4 \) are essentially flat against the firm’s sales growth between quarters \( t - 1 \) and \( t \) – the quarter just prior to the forecast. By contrast, realized sales growth between \( t \) and \( t + 4 \) correlates negatively with the firm’s lagged performance. This pattern suggests managers forecasts fail to internalize that the current shock decays over time, thus making their errors predictable because this decay is predictable. In this sense, managers overextrapolate from the level of current sales rather than the recent rate of sales growth. This finding is consistent with how I model overextrapolation in Section 3, in which firms receive shocks that shift the level of profitability in a stationary environment.

#### A.8.2 Forecast errors are negatively correlated with past forecast errors

Figure A.21 is a bin-scatter plot of forecast errors on lagged forecast errors, showing a clear negative relationship. The horizontal axis plots twenty quantiles of forecast minus realized sales growth for quarters \( t - 4 \) to \( t \) against the mean forecast minus realized sales growth for \( t \) to \( t + 4 \) on the vertical axis. Managers that fall on the right half of the graph are those who made forecasts on \( t - 4 \) that ended up overestimating the firm’s actual sales growth between \( t - 4 \) and \( t \). Those same managers then subsequently make forecasts on \( t \) for sales growth between \( t \) and \( t + 4 \) and end up underestimating. This pattern is consistent with my finding in Section 2 that managers overextrapolate. Namely, those who receive a negative shock between \( t - 4 \) and \( t \) perceive that negative shock to be particularly persistent and thus end up underestimating as they look forward from \( t \) to \( t + 4 \).

In Table A.3 I show estimates from the regression depicted in Figure A.21, as well as from specifications that add date, sector-by-date, and firm fixed effects. The existence of a negative relationship is highly robust across all specifications, although the coefficient from the fourth column with firm and date fixed effects is significantly larger. This is probably due to the fact that I have a short of the panel in which fixed effects specifications may be upward biased. (Recall that I only have data since late 2014 and I need two years...
worth of data of an individual firm to obtain an observation with non-missing current and lagged forecast errors.)

A.8.3 Forecast errors are positively correlated with a second measure of sales growth

Recall that in the SBU questionnaire managers report their firm’s sales growth in the past twelve months (see Figure 2 in the main text). Figure A.22 shows that managers’ forecast errors in the SBU are positively related with their reported sales growth for the twelve months prior to. Again, I interpret this figure as evidence of over extrapolation. Managers at firms that receive one or more positive shocks in the year up to \( t \) overestimate the persistence of those shocks and thus their forecasts for sales growth between \( t \) and \( t+4 \) overestimate the firm’s performance looking forward. Managers at firms that receive one or more negative shocks in the past year, in turn, are too pessimistic about future sales growth.

Table A.4 shows that the relationship between reported past sales growth and forecast errors is robust to controlling for date, sector-by-date, and firm fixed effects. The final column that includes both firm and date fixed effects once again has a larger coefficient, potentially due to biases that are common in dynamic panel regression models. However, the coefficient does not look statistically different from that in the first three columns, so the overall picture is that the relationship between reported past sales growth and forecast errors is robust.

A.9 Managerial Optimism, Overconfidence, and Overextrapolation about Future Employment

Although I focus on biases in managers’ beliefs about their firm’s future sales growth, the SBU also collects subjective probabilities and tracks outcomes for the firm’s level of employment. See Figure A.11 for the SBU’s questions about current and future employment. Here, I document as a robustness check that managers do not appear to be particularly over-optimistic nor pessimistic about their firms’ future employment growth, but they do appear to be overconfident and overextrapolate.

Table A.5 summarizes the results about managers’ optimism/pessimism, overconfidence, and overconfidence about their firm’s future employment growth. In Panel A, we can see that managers’ appear somewhat pessimistic about their firm’s future employment growth on average. Realized employment growth exceeds its ex-ante forecast by about 0.016 on average, and is statistically significantly different from zero with 95 percent but not 99 percent confidence. Economically speaking, however, this is arguably not a significant deviation from rational expectations, especially given that the standard deviation of realized employment growth over four quarters is 0.175.

Panel B of Table A.5 shows that managers are overconfident about future employment. While the mean absolute forecast error for employment growth is close to 0.11, under managers’ subjective distributions the mean absolute error should only be less than half as big at 0.044. This means there is an excess absolute error of about 0.066.

Turning now to Panel C, we can see that forecast minus realized sales growth for months \( t \) to \( t+12 \) correlates significantly with firms’ employment growth in months \( t-2 \) to \( t \), just prior to managers’ making their forecasts. As with sales growth, the positive relationship between lagged employment growth and forecast errors means that managers at firms performing well when they answer the survey end up overestimating future employment growth, and vice-versa for those at firms performing poorly. I focus here on twelve-month changes in future employment because this is the horizon used for the employment questions in Figure A.11. Similarly I use changes in employment from \( t-2 \) to \( t \) as an independent variable because respondents get the questionnaire for employment every two months. Neither of these particular choices are crucial to the results. In the second and third columns I include date and firm fixed effects to show that the predictability of forecast errors is not driven by aggregate shocks or by persistent differences across firms.

A.10 Why don’t market forces throw out biased managers?

If managers are truly biased and make systematic mistakes that destroy firm value, we might intuitively expect boards of directors and headhunting firms should realize this and prevent these biased individuals
from taking on managerial positions. In practice, it is not obvious that these other market participants can easily gather the data required to make such assessments.

First, individual point forecasts may be at odds with realizations due to random shocks even if managers have rational expectations. To determine whether an individual manager is biased, you would need to show his or her ex-ante beliefs are *systematically* inconsistent with realizations, which requires several observed forecasts and realizations. Since firm outcomes are typically reported at relatively low frequencies, namely annually or quarterly, it may take years or decades to obtain enough statistical power to determine if, say, managers’ forecasts are over-optimistic on average.

In my survey data, the volatility of sales growth is such that even after tracking forecasts and realizations for 25 years at a quarterly frequency (yielding 100 observations), one would typically not be able to determine with 95 percent confidence whether a manager who over-estimates her firm’s sales growth by as much as 5 percentage points on average is truly over-optimistic. Given the median CEO tenure is only about 7 years (e.g. see Taylor, 2010), boards typically have a fraction of the data and statistical power I used in this example. Similarly, this small-sample problem may preclude individual managers from figuring out whether their beliefs about a given firm’s risks and prospects are biased or not. By the time they have enough data, they may move to a different firm with potentially different conditions and may thus have to start learning about its profitability all over again.

Second, determining whether managers are biased requires reliable data on their beliefs. If, say, a board of directors routinely asks managers for their own subjective forecasts of the firms’ future performance, it is hard to think how the board could compel them to report their beliefs honestly. For example, the board may not be able to commit to not using those forecasts to judge the managers’ strategy, thus incentivizing managers to misreport. Note this is in contrast with my survey data, in which individual responses are confidential and never reported individually.

Finally, if a board somehow establishes that it will only use managerial forecasts to determine whether managers are biased, this may be an incentive for managers to report honest forecasts, but it could also create incentives for managers to manipulate the firm’s reported or real performance ex-post. Thus, the board may rationally choose to refrain from systematically tracking managerial forecasts accuracy to avoid introducing these additional incentives for managerial misbehavior.
Notes: Sales growth questions in the *Survey of Business Executives* as they appeared prior to September 2016. In months prior to September 2016, the SBE asked for sales growth beliefs in levels rather than growth rates. See Appendix Figure
Figure A.2: Forecast Error Observations from the SBU Belong to Firms From All Sector

Notes: Number of forecast error observations by one-digit sector. Data are from the SBU covering the period 10/2014 to 6/2018, restricting attention to subjective probability distributions. A forecast error observation consists of a response in quarter $t$ with a well-formed subjective probability distribution for sales growth, looking 4 quarters ahead, for which I also observe realized sales growth between quarters $t$ and $t+4$. $N = 1,574$.

Figure A.3: SBU Respondents Are Larger, Well-Established Firms (Employment)

Notes: Distribution of current employment (winsorized at the top and bottom 5-percent) at the time of forecast for all forecast error observations in the Survey of Business Uncertainty. The sample period includes all months between 10/2014 to 6/2018. A forecast error observation consists of a response in quarter $t$ with a well-formed subjective probability distribution for sales growth, looking 4 quarters ahead, for which I also observe realized sales growth between quarters $t$ and $t+4$. $N = 1,574$. 
Figure A.4: **SBU Respondents Are Larger, Well-Established Firms (Sales)**

![Figure A.4: SBU Respondents Are Larger, Well-Established Firms (Sales)](image)

**Notes:** Distribution of current sales ($M$, winsorized at the top and bottom 5-percent) at the time of forecast for all forecast error observations in the *Survey of Business Uncertainty*. The sample period includes all months between 10/2014 to 6/2018. A forecast error observation consists of a response in quarter $t$ with a well-formed subjective probability distribution for sales growth, looking 4 quarters ahead, for which I also observe realized sales growth between quarters $t$ and $t+4$. $N = 1,574$.

Figure A.5: **Firms with Forecast Errors are Older**

![Figure A.5: Firms with Forecast Errors are Older](image)

**Notes:** Number of forecast error observations, sorting firms by the decade in which they hired their first paid employee. The sample period includes all months between 10/2014 to 6/2018. A forecast error observation consists of a response in quarter $t$ with a well-formed subjective probability distribution for sales growth, looking 4 quarters ahead, for which I also observe realized sales growth between quarters $t$ and $t+4$. Data about when the firm hired its first paid employee comes from a one-off special question included in the January 2017 survey, asking "In what year did your firm hire its first paid employee? If you do not know the precise year please give your best estimate." The respondent then was able to select either an individual year since 2000, or one of "1990-1999", "1980-1989", "1970-1979", or "prior to 1970". $N = 1,414$. 

13
Notes: Distribution of four-quarter sales growth realizations (winsorized at the top and bottom 1 percent) for all forecast error observations in the *Survey of Business Uncertainty*. The sample period includes all months between 10/2014 to 6/2018. A forecast error observation consists of a response in quarter \( t \) with a well-formed subjective probability distribution for sales growth, looking 4 quarters ahead, for which I also observe realized sales growth between quarters \( t \) and \( t + 4 \). \( N = 1,574 \).

Figure A.7: **Share of Employment by Firm Size: SBU vs. US Economy**

Notes: This figure shows (1) the share of employment across all SBU responses from 10/2014 to 7/2018 made by firms in each firm size category; (2) the share of employment for each firm size category in the US economy in the US Census Bureau’s 2015 Statistics on US Businesses.
**Figure A.8: Share of Employment by Firm Age: SBU vs. US Economy**

![Firms by Year of Birth Diagram](image)

**Notes:** This figure shows (1) the share of employment across all SBU responses from 10/2014 to 7/2018 by the firm’s year of birth; (2) the share of employment across firms by year of birth in the US economy according to the US Census Bureau’s 2015 Business Dynamics Statistics.

**Figure A.9: Share of Employment by Sector: SBU vs. US Economy**

![Employment by Sector Diagram](image)

**Notes:** This figure shows (1) the share of employment across all SBU responses from 10/2014 to 7/2018 made by firms in each sector; (2) the share of employment in each sector of the US economy in the US Census Bureau’s 2015 Statistics on US Businesses. Numbers in parentheses correspond to NAICS two-digit codes for each sector.
Figure A.10: **Share of Employment by Region: SBU vs. US Economy**

**Notes:** This figure shows (1) the share of employment across all SBU responses from 10/2014 to 7/2018 made by firms in each region (i.e. Census Division); (2) the share of employment in each region according the US Census Bureau’s 2015 Statistics on US Businesses.
Figure A.11: SBU Questions About Employment

Notes: This figure shows the questions about current employment and beliefs about future employment in the Survey of Business Uncertainty.
Figure A.12: Sales and Employment Growth Forecasts Predict Outcomes

(a) Sales

(b) Employment

Notes: This figure shows bin-scatter plots of sales growth and employment growth forecasts on the horizontal axis against realized sales and employment growth. Sales growth forecasts are made in quarter $t$ and forecast sales growth between quarter $t$ and $t + 4$. Employment growth forecasts are made in month $m$ and forecast employment growth between $m$ and $m + 12$. T-statistics for the underlying regressions are 4.9 and 7.7 for sales and employment, respectively using firm-clustered standard errors. All data are from the SBU with the sample period covering 10/2014 to 6/2018. The plot for sales includes 1,574 forecast error observations from 408 firms, and the employment plot includes 2,143 observations from 460 firms.
Figure A.13: Planned Hiring vs. Sales Growth Forecasts

Notes: This figure shows a bin-scatter plot of planned net hiring (i.e. expectations for employment growth) looking forward a year against the managers' sales growth forecast looking ahead to the next four quarters. All data come from the SBU with the sample period covering 10/2014 to 6/2018. The underlying regression for the figures above includes 3,615 SBU responses from 695 firms.

Figure A.14: Actual Hiring vs. Actual Sales Growth

Notes: This figure shows a bin-scatter plot of actual net hiring (i.e. employment growth) in the year following a forecast against the actual sales growth recorded in the four quarters following the forecast. All data come from the SBU with the sample period covering 10/2014 to 6/2018. The underlying regression for the figures above includes 1,234 SBU responses from 330 firms.
Figure A.15: **Actual Hiring vs. Sales Growth Forecasts**

**Notes:** This figure shows a bin-scatter plot of actual net hiring (i.e. employment growth) in the year following a forecast against the managers’ sales growth forecast looking ahead over four quarters. All data come from the SBU with the sample period covering 10/2014 to 6/2018. The underlying regression for the figures above includes 1,514 SBU responses from 365 firms.
Figure A.16: **Current Hiring vs. Sales Growth Expectations and Uncertainty**

(a)

![Graph showing current hiring vs. sales growth forecasts for quarters t to t+4](image)

Net Hiring Quarter t

Forecast Sales Growth, Quarter t \(\text{to}\) t+4

\(b = 0.102 \ (0.054)\), \(N = 2834\)

(b) **Current Hiring vs. Recent Growth**

![Graph showing current hiring vs. recent sales growth](image)

Net Hiring t

Sales Growth Quarter t-1 \(\text{to}\) t

\(b = 0.068 \ (0.016)\), \(N = 2695\)

**Notes:** This figure shows bin-scatter plots of current hiring in quarter \(t\) against managers’ sales growth forecasts for quarters \(t\) to \(t + 4\) (top) and the firm’s sales growth between quarters \(t - 1\) and \(t\). All data come from the SBU with the sample period covering 10/2014 to 6/2018.
Figure A.17: Managers are Overconfident About Their Forecasts’ Accuracy

Notes: This figure plots the empirical distribution of absolute forecast errors as well as the distribution of absolute errors that would arise if sales growth realizations were drawn from SBU respondents’ subjective probability distributions. I scale each distribution so that the sum of the heights of the bars is equal to one, and fix the width of the bars to 0.05. The sample period includes all months between 10/2014 to 6/2018. A forecast error observation consists of a response in quarter $t$ with a well-formed subjective probability distribution for sales growth, looking 4 quarters ahead, for which I also observe realized sales growth between quarters $t$ and $t + 4$. $N = 1,574.$
Figure A.18: Managers are Overconfident Across Levels of Subjective Uncertainty

Notes: Bin-scatter plot of realized and subjective absolute forecast errors for sales growth looking four quarters ahead, against ex-ante subjective uncertainty (the standard deviation of respondents' subjective distribution). A respondent’s subjective absolute forecast error is the subjective mean absolute deviation from her forecast. Data are from the SBU and sample period includes all months between 10/2014 to 6/2018. A forecast error observation consists of a response in quarter $t$ with a well-formed subjective probability distribution for sales growth, looking 4 quarters ahead, for which I also observe realized sales growth between quarters $t$ and $t + 4$. $N = 1,574$. 
Figure A.19: Forecast Errors in the SBU vs. IBES

(a) Empirical Distributions of Forecast Errors in SBU & IBES

(b) Subjective Distribution of Forecast Errors in the SBU vs. Empirical Distribution in IBES

Notes: The top figure shows; (1) the empirical distribution of managers’ forecast errors for sales growth looking four quarters ahead from the SBU; (2) the empirical distribution of analyst forecast errors for sales growth four quarters ahead from IBES. The bottom figure shows (1) the subjective distribution of managers’ forecast errors for sales growth looking four quarters ahead from the SBU (i.e. the distribution of forecast errors implied by managers’ subjective probabilities); (2) the empirical distribution of analyst forecast errors for sales growth four quarters ahead from IBES. The SBU sample includes 1,574 forecast error observations from 397 firms between 10/2014 and 6/2018. The IBES sample includes 755,685 analyst forecast errors.
Figure A.20: **Overextrapolation Arises Because Managers Ignore Mean Reversion**

![Graph: Overextrapolation Arises Because Managers Ignore Mean Reversion](image)

**Notes:** This figure shows bin-scatters of forecast and realized sales growth between \( t \) and \( t + 4 \) on the vertical axis against realized sales growth between quarters \( t - 1 \) and \( t \), just prior to the survey response. Data are from the SBU and sample period includes all months between 10/2014 to 6/2018. A forecast error observation consists of a response in quarter \( t \) with a well-formed subjective probability distribution for sales growth, looking 4 quarters ahead, for which I also observe realized sales growth between quarters \( t \) and \( t + 4 \). \( N = 919 \).

Figure A.21: **Managers Overextrapolate: Forecast Errors Serially Correlated**

![Graph: Managers Overextrapolate: Forecast Errors Serially Correlated](image)

**Notes:** This figure shows a bin-scatter plot of forecast minus realized sales growth over quarters \( t \) to \( t + 4 \) on the vertical axis against forecast minus realized sales growth over quarters \( t - 4 \) to \( t \). Data are from the SBU covering 10/2014 to 6/2018. \( N = 502 \).
Figure A.22: Managers Overextrapolate: Based on Reported Sales Growth

Notes: This figure shows a bin-scatter plot of forecast minus realized sales growth over quarters $t$ to $t + 4$ on the vertical axis against the managers’ reported sales growth for quarters $t - 4$ to $t$. Data are from the SBU covering 10/2014 to 6/2018. $N = 1,071$. 

26
Table A.1: **SBU Sample Descriptives: Firms with Forecast Errors**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Median</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Employment</td>
<td>477.1</td>
<td>175</td>
<td>1823</td>
</tr>
<tr>
<td>Quarterly Sales ($M)</td>
<td>38.1</td>
<td>8.00</td>
<td>232</td>
</tr>
<tr>
<td>4-Qtr Sales Growth</td>
<td>0.046</td>
<td>0.043</td>
<td>0.266</td>
</tr>
</tbody>
</table>

**Notes:** Summary statistics on current employment, current sales, and measured sales growth realizations among all forecast error observations in the *SBU*. An observation here is a survey response in quarter $t$ with a well-defined subjective probability distribution for sales growth looking four quarters ahead, and for which I also observe the firm’s realized sales growth between quarters $t$ and $t+4$.

Table A.2: **Discretizing Empirical Distributions**

(a) "Tauchen" (Equidistant-Bins) Approach

<table>
<thead>
<tr>
<th>Mass Excluded ($p$)</th>
<th>0.01</th>
<th>0.05</th>
<th>0.1</th>
<th>0.2</th>
<th>0.4</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Excess Absolute Fcast. Error</td>
<td>0.024</td>
<td>0.012</td>
<td>0.022</td>
<td>0.042</td>
<td>0.075</td>
<td>0.145</td>
</tr>
</tbody>
</table>

(b) Quantile Approach

<table>
<thead>
<tr>
<th>Mass Excluded ($p$)</th>
<th>0.01</th>
<th>0.05</th>
<th>0.1</th>
<th>0.2</th>
<th>0.4</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Excess Absolute Fcast. Error</td>
<td>-0.01</td>
<td>0.017</td>
<td>0.031</td>
<td>0.045</td>
<td>0.058</td>
<td>0.145</td>
</tr>
</tbody>
</table>

**Notes:** The above tables show the excess absolute forecast error that would arise from approximating the empirical distribution of realized sales growth between quarters $t$ and $t+4$ under the "Tauchen"-based, and Quantile-based approaches to discretization. Before discretizing, I remove heterogeneity in realized sales growth attributable to differences in subjective first moments, leaving the empirical distribution of realized sales growth for the typical expectation and subjective uncertainty across all 1,574 forecast error observations in the SBU. See Appendix A.7 for a full description of the two discretization approaches.
Table A.3: Managers Overextrapolate: Forecast Errors are Serially Correlated

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forecast - Realized Sales Growth, quarters t-4 to t</td>
<td>-0.201***</td>
<td>-0.212***</td>
<td>-0.161**</td>
<td>-0.508***</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.020</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Date FE</td>
<td></td>
<td>Y</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Date x Sector FE</td>
<td></td>
<td></td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Firm FE</td>
<td></td>
<td></td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>502</td>
<td>502</td>
<td>428</td>
<td>451</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.042</td>
<td>0.068</td>
<td>0.241</td>
<td>0.495</td>
</tr>
</tbody>
</table>

Notes: Robust standard errors in parentheses, clustered by firm. Data are subjective probability distributions and forecast errors about sales growth looking 4 quarters ahead from the Survey of Business Uncertainty covering all months between October 2014 and June 2018. An observation is a forecast error for a particular firm in a particular month. *** p<0.01, ** p<0.05, * p<0.1

Table A.4: Managers Overextrapolate: Based on Reported Sales Growth

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forecast - Realized Sales Growth, quarters t to t+4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reported Sales Growth, 12 months up to t</td>
<td>0.261**</td>
<td>0.267***</td>
<td>0.266**</td>
<td>0.390***</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.018*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Date FE</td>
<td>Y</td>
<td></td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Date x Sector FE</td>
<td>Y</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Firm FE</td>
<td>Y</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>1,071</td>
<td>1,071</td>
<td>1,062</td>
<td>1,021</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.021</td>
<td>0.068</td>
<td>0.241</td>
<td>0.495</td>
</tr>
</tbody>
</table>

Notes: Robust standard errors in parentheses, clustered by firm. Data are subjective probability distributions and forecast errors about sales growth looking 4 quarters ahead from the Survey of Business Executives covering all months between October 2014 and February 2018. An observation is a forecast error for a particular firm in a particular month. *** p<0.01, ** p<0.05, * p<0.1
Table A.5: Beliefs Biases About Future Employment Growth

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A. Optimism</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Employment Growth Forecast Error</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>0.009</td>
<td>0.025</td>
<td>-0.016</td>
</tr>
<tr>
<td>SE</td>
<td>(0.004)</td>
<td>(0.006)</td>
<td>(0.005)</td>
</tr>
<tr>
<td>Obs.</td>
<td>2,143</td>
<td>2,143</td>
<td>2,143</td>
</tr>
<tr>
<td>Firms</td>
<td>460</td>
<td>460</td>
<td>460</td>
</tr>
<tr>
<td><strong>Panel B. Overconfidence</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Absolute Forecast Error Excess Error</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>0.111</td>
<td>0.044</td>
<td>0.066</td>
</tr>
<tr>
<td>SE</td>
<td>(0.005)</td>
<td>(0.003)</td>
<td>(0.004)</td>
</tr>
<tr>
<td>Obs.</td>
<td>2,143</td>
<td>2,143</td>
<td>2,143</td>
</tr>
<tr>
<td>Firms</td>
<td>460</td>
<td>460</td>
<td>460</td>
</tr>
<tr>
<td><strong>Panel C. Overextrapolation</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dependent Variable</td>
<td>Forecast - Realized Emp. Growth, months $t$ to $t + 12$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Emp. Growth, months. $t - 2$ to $t$</td>
<td>0.356</td>
<td>0.357</td>
<td>0.470</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.026</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Date FE</td>
<td>Y</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Firm FE</td>
<td>Y</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Obs.</td>
<td>1,088</td>
<td>1,088</td>
<td>1,035</td>
</tr>
<tr>
<td>Firms</td>
<td>299</td>
<td>299</td>
<td>246</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.070</td>
<td>0.071</td>
<td>0.545</td>
</tr>
</tbody>
</table>
B Model Appendix

B.1 An Alternative Way of Writing Down Managers’ Optimization Problem and True Firm Value

Equations 4 and 5 in the main text express managers’ problem and the firm’s true value recursively. Although I use these recursive forms to compute managers’ policy functions, subjective valuations of the firm, and the firm’s true value, the same objects can be represented as a sequence problem that may be more intuitive to some readers.

Since managers are risk-neutral and own a share $\theta \in (0, 1]$ of the firm, their objective is to maximize the net present value of the firm’s cash flows, discounted at the equilibrium risk-free rate. Managers take as given the full path of current and future risk-free rates $\{r_{t+1}\}_{t=0}^{\infty}$ and wages $\{w_{t}\}_{t=0}^{\infty}$, which are deterministic.

However, they forecast their firms’ future business conditions $z_{t+k}$ for $k > 0$ under their own subjective beliefs. The following sequence problem thus represents managers’ problem, given some initial business conditions $z_0$ and labor $n_0$:

$$\tilde{V}(z_0, n_0) = \max_{\{n_{t+1}\}_{t=0}^{\infty}} \tilde{E}_0 \left[ \sum_{t=0}^{\infty} \pi(z_t, n_t, n_{t+1}; w_t) \frac{1}{R_t} \right]$$

where $R_t$ is the (non-stochastic) composite discount factor used to value period-$t$ cash flows in period 0:

$$R_t = \Pi_{s=0}^{t-1} (1 + r_{s+1})$$

and, again, operator $\tilde{E}[\cdot]$ denotes the managers’ subjective expectations based on her stochastic process from equation 3 in the main text. Looking at this formulation, it is clear that the share of the firm actually owned by the manager $\theta$ is irrelevant for their optimal policy.

From standard dynamic programming results (e.g. see Stokey et al., 1989) we know that managers’ hiring policy on date $t$ can be written solely as a function of current business conditions and labor: $n_{t+1} = \kappa(z_t, n_t)$. Taking as given a manager’s policy function $\kappa(\cdot)$, the objective, expected net present value of the firm’s cash flows is:

$$V(z_0, n_0) = E_0 \left[ \sum_{t=0}^{\infty} \pi(z_t, n_t, \kappa(z_t, n_t); w_t) \frac{1}{R_t} \right]$$

where I have substituted in the manager’s policy. In contrast with the managers’ valuation, computing the true value of the firm requires forecasting future cash flows based on the unbiased expectation $E[\cdot]$.

Implicit in both the managers’ subjective and the true value of the firm is the capacity to forecast future prices $w_t$ and $R_t$ perfectly. This in practice follows from my assumption that there is no aggregate risk both in the baseline model from Section 3 and the investment-based model in B.5. Then, an equilibrium includes a deterministic sequence of prices that I assume model agents know with certainty.

B.2 Definition of Aggregate Quantities

Here I define aggregate quantities in my model economy from Section 3 in the main text. Below I use $\Phi(z, n)$ to denote the measure of firms in the economy with business conditions $z$ and labor $n$.

Aggregate output or GDP in my model economy is the sum of value added across all firms less spending.

41 I use the dynamic programming result here for simplicity of exposition. In practice, we could denote $\{n_{t+1}\}_{t=0}^{\infty}$ to be the manager’s optimal policy and let it be a function of the entire history of shocks and expectations of future shocks.
on adjustment costs:

\[ Y = \int_{z \times N} zn^\alpha - \lambda \left( \frac{\kappa(z, n) - (1 - q)n}{n} \right)^2 n \, d\Phi(z, n) \]

\[ = \hat{Y} - AC. \]

Here I use \( \hat{Y} \) to denote gross output (before subtracting adjustment costs) and \( AC \) total spending on adjustment costs. This definition of GDP is crucial for parts of my analysis about the aggregate implications of beliefs biases in Section 5.2 of the main text. I justify subtracting adjustment costs from GDP because resources spent in this way do not constitute income for any agents in the economy and instead are essentially intermediate business expenses that subtract from profits and thus value added.

Recall that managers in the model are risk neutral and own a share \( \theta = \in (0, 1] \). They consume \( \theta \) times the firm’s current cash flow \( \pi(\cdot) \), while the rest of the firm’s cash flow goes to the representative household. As stated in the main text, the household then receives capital income \( \theta \Pi \) where

\[ \Pi = \int_{z \times N} \pi(z, n, \kappa(z, n); w) \, d\Phi(z, n) \]

and \( \kappa(z, n) \) is the hiring policy of a manager at a firm with state \( (z, n) \).

It follows that aggregate output must be equal to the household’s consumption plus the managers’ consumption, or equivalently the sum of labor and capital income:

\[ Y = C + \theta \Pi \]

\[ = wN + \Pi \]

**B.3 Firm Value Welfare Change and Formulas**

To compute the change in firm value from replacing a biased manager for a counterfactual manager who is unbiased (or less biased), I first compute each of their policy functions \( \kappa(\cdot) \) and \( \kappa^c(\cdot) \) respectively. Based on those policy functions, I find the true net present value of cash flows generated by their respective policies, \( V(\cdot) \) and \( V^c(\cdot) \), by solving the functional equation in 5 in the main text. See Appendix C.3 for more details about computing true firm value. The average percent change in firm value obtained from replacing a biased manager is then:

\[ \mathbb{E}[\Delta V] \% = 100 \cdot \int_{z \times N} \left[ \frac{V^c(z, n)}{V(z, n)} - 1 \right] \, d\Phi(z, n). \]

The consumption-equivalent difference in welfare between my baseline economy with consumption \( C \) and aggregate labor \( N \) and a counterfactual economy with \( C_c \) and \( N_c \), both in their long-run stationary general equilibrium is \( 100 \times \xi \), where \( \xi \) satisfies:

\[ \sum_{t=0}^{\infty} \beta^t \left[ \frac{C(1 + \xi)^{1-\gamma}}{1 - \gamma} - \chi \frac{N^{1+\eta}}{1+\eta} \right] = \sum_{t=0}^{\infty} \beta^t \left[ \frac{C_c^{1-\gamma}}{1 - \gamma} - \chi \frac{N_c^{1+\eta}}{1+\eta} \right] \]

which has a simple closed form solution:

\[ \xi = \left[ \frac{C_c^{1-\gamma} - \chi^{1-\gamma} \frac{N_c^{1+\eta} - N^{1+\eta}}{1+\eta}}{C} \right]^{1-\gamma} - 1. \]
B.4 Capital and Intermediate Goods in my Baseline Model

My baseline setup in Section 3 assumes sales are a function of just labor. By omitting capital, I’m implicitly assuming the firm has a fixed stock of physical and intangible capital and that changes in that capital stock are part of the firm’s profitability shock. Given capital moves relatively slowly and my model has a quarterly frequency this assumption seems fairly reasonable.

My sales function in Section 3 also abstract from intermediate goods that may be used in product. Implicitly, you could imagine that the firm consumes some of the final good in the economy as an intermediate in production. To the extent the underlying production function for gross output has the same Cobb-Douglas assumption I assume in my baseline setup, including intermediates that the firm optimizes statically, then you would get a value-added production function of the form I assume in Section 3.

B.5 Investment Model

B.5.1 Technology

As in Section 3, there is a continuum of firms with access to a decreasing returns to scale revenue production function using capital \( k \), labor \( n \) and a Hick-neutral shock \( \hat{z} \):

\[
\hat{y}(\hat{z}_t, k_t, n_t) = \hat{A}\hat{z}_t^{\hat{\alpha}} k_t^{\hat{\mu}} n_t^{\hat{\nu}}.
\]

The constant \( \hat{A} \) is a scaling factor.

The firm hires labor statically in a spot market and pays the equilibrium wage \( w_t \). After making its optimal labor choice, the firm receives earnings or operating income (= revenue minus wage bill) that depends on the capital stock it had entering the period and the equilibrium wage:

\[
y(z_t, k_t; w_t) = A(w_t)z_t^{\alpha} k_t^{\alpha} n_t^{\alpha} (1-\gamma)
\]

where \( y(\cdot) \) now denotes earnings as opposed to revenue, \( z_t = \hat{z}_t^{1/(1-\gamma)} \) is a renormalization of the firm’s shock, and \( \alpha = \hat{\alpha} / (1 - \hat{\nu}) \). The function \( A(\cdot) \) is decreasing, so higher wages result in lower earnings for the same amount of capital.

The firm’s capital stock follows a standard law of motion, so that the amount of capital in place for quarter \( t+1 \) depends on the previous quarter’s capital less depreciation plus investment:

\[
k_{t+1} = (1-\delta)k_t + i_t.
\]

Investment is subject to adjustment costs, which I assume to have a two components: (1) smooth quadratic costs in the gross investment rate, and (2) partial irreversibility that means re-sold capital is sold at a discount. The total cost of obtaining capital \( k_{t+1} \) next quarter starting from capital \( k_t \) is then:

\[
AC(k_t, k_{t+1}) = \begin{cases} 
\frac{k_{t+1} - (1-\delta)k_t}{k_t} \\
\lambda_i [k_{t+1} - k_t(1-\delta)] \cdot 1(k_{t+1} < k_t(1-\delta)) + \lambda_q k_t \left( \frac{k_{t+1} - (1-\delta)k_t}{k_t} \right)^2 
\end{cases}
\]

The first term represents the cost of acquiring new capital, the second implies a loss of \( \lambda_i \in [0,1] \) for any capital that is re-sold, and \( \lambda_q \) scales the firm’s quadratic adjustment costs. Following the consensus in the literature (e.g. see Cooper and Haltiwanger (2006), Bloom (2009), and Winberry (2015)) that capital adjustment typically involves both convex and non-convex costs, I include one type of each. As in the main, labor-based model from Section 3, the magnitude and form of adjustment costs matters for the quantitative implications of biases in my model.

The firm’s cash flow in period \( t \) equals earnings (itself revenue minus labor costs) less capital adjustment.
costs:
\[ \pi(z_t, k_t, k_{t+1}; w_t) = y(z_t, k_t; w_t) - AC(k_t, k_{t+1}) \]

### B.5.2 Subjective and Objective Shock Processes

As before, shocks to the firm’s earnings follow a Gaussian autoregressive process in logs:
\[ \log(z_{t+1}) = \mu + \rho \log(z_t) + \sigma \varepsilon_{t+1}, \quad \varepsilon_{t+1} \sim \mathcal{N}(0, 1) \]
but the manager running the firm may have incorrect beliefs about the mean, persistence, and standard deviation of innovations to \( \log(z_t) \). She believes:
\[ \log(z_{t+1}) = \tilde{\mu} + \tilde{\rho} \log(z_t) + \tilde{\sigma} \varepsilon_{t+1}. \]

### B.5.3 Manager’s Decision Problem and True Firm Value

Also as in the baseline model from Section 3, I assume each manager running a firm in the economy is compensated with a share \( \theta \in (0, 1) \) of her firm’s equity and therefore aims to maximize the net present value of her firm’s cash flows (discounted at the equilibrium risk-free rate). This optimization requires the manager to forecast future shocks \( z_t \), which she does under her subjective beliefs. The manager’s recursive problem is therefore:
\[
\tilde{V}(z_t, k_t; w_t, r_{t+1}) = \max_{k_t > 0} \pi(z_t, k_t, k_{t+1}; w_t) + \frac{1}{1 + r_{t+1}} \tilde{E}[\tilde{V}(z_{t+1}, k_{t+1}; w_{t+1}, r_{t+2})]
\]  

(10)

where \( \tilde{E}[\cdot] \) denotes the manager’s expectations under her subjective beliefs.

Given a manager’s policy \( k_{t+1} = \kappa(z_t, k_t; w_t, r_{t+1}) \), the true net present value of the firm’s cash flows \( V(\cdot) \) (without a tilde) is:
\[
V(z_t, k_t; w_t, r_{t+1}) = \pi(z_t, k_t, \kappa(z_t, k_t); w_t) + \frac{1}{1 + r_{t+1}} E[V(z_{t+1}, \kappa(z_t, k_t); w_{t+1}; r_{t+2})]
\]

where \( E[\cdot] \) takes expectations according to the objective stochastic process for \( z_{t+1} \).

### B.5.4 Household and Equilibrium

Again, there is an aggregate household that consumes, supplies labor, and owns the remaining \((1 - \theta)\) equity in the firms in the economy. The household’s problem is identical to that of Section 3.5.

A stationary equilibrium in this specification of the model is then a set of prices \( \{w, r\} \), consumption, labor supply and saving choices by the household \( C, N^S, B \), subjective firm valuations \( \tilde{V}(z_t, k_t; w_t, r) \) made by managers, and a stationary distribution \( \phi : Z \times \mathcal{N} \to [0, 1] \) such that:

1. \( \tilde{V}(z_t, k_t; w, r) \) solves each managers’ problem in 10.

2. The household optimally chooses steady-state consumption \( C \), labor supply \( N^S \), and savings \( B = 0 \) (in zero-net-supply by assumption).

3. The distribution of firms \( \phi(\cdot) \) is invariant across quarters and consistent with managers’ hiring and firing decisions and exogenous fluctuations in business conditions of incumbents:
\[
\phi(z', k') = \int_{Z \times K} \phi(z, k) \cdot Pr(z'|z) \cdot 1(k' = \kappa(z, k; w, r))dzdk
\]
4. The labor market clears:

\[ NS = ND = \int_{z,k} n(z,k;w)\phi(z,k)dzd \]

C Simulation Appendix

C.1 Model Solution Details

Here I provide some additional details regarding the algorithm I use to solve for managers’ dynamic hiring problem in equation 4 of the main text. I solve for managers’ value and policy functions over a discretized \((z,n)\) state space employing policy function iteration. I choose grids of size \((21,100)\) since the managers’ dynamic program is standard and, by contemporary standards, not computationally intensive with only two state variables. As is standard for numerical dynamic programming I make my grid for possible labor choices \(n\) linear in log-space and make the end-points of the grid far out enough so that under the stationary distribution of the estimated model \(\phi(z,n)\) there is near zero probability of ending up in the highest and lowest grid points, i.e. so \(\max_z \{\phi(z,n_1), \phi(z,n_{100})\} < 10^{-5}\).

C.1.1 Discretizing the subjective and objective driving processes

I approximate both the objective and subjective stochastic processes for \(\log(z_t)\) on a single set of grid points using the algorithm of Tauchen (1986). My choice of 21 grid points for \(\log(z)\) is dense enough to achieve an accurate approximation of the conditional first and second moments of both the subjective and objective AR(1) processes on the same grid. Representing the two stochastic processes on the same discrete grid is computationally convenient and seems a reasonable choice so that managers are only wrong about the probability of a given event happening, but they are correct about the set of event that may potentially happen. In practice, I pick the set of potential grid points to be symmetric around zero (given both the true and subjective stochastic processes are approximately mean zero in the model). For any given set of \(\sigma, \tilde{\sigma}, \rho, \text{ and } \tilde{\rho}\), the highest and lowest grid points for \(\log(z_t)\) are at \(\pm 2.575 \sqrt{\hat{\sigma}^2 \frac{1}{1-\rho^2}}\) where \(\hat{\sigma} = \max\{\sigma, \tilde{\sigma}\}\) and \(\hat{\rho} = \max\{\rho, \tilde{\rho}\}\) so that the grid covers 99 percent of the support of a Gaussian AR(1) process with mean zero and the largest unconditional standard deviation possible given the parameters fed into the model.

C.1.2 Computing managers’ optimal policies and subjective firm valuations

Given some value of the stationary equilibrium wage \(w\), I solve for managers’ optimal subjective valuation of the business in 4 numerically using standard techniques. Specifically, I solve for managers’ value and policy functions over a discretized \((z,n)\) state space via policy function iteration (i.e. value function iteration aided by Howard’s improvement algorithm). The only noteworthy detail for this procedure is I use managers’ subjective beliefs for the evolution of \(z_t\) (instead of the true stochastic process) to forecast managers’ expectation of the firm’s future (subjective) value.

Starting with a guess for the managers’ subjective valuation of the firm \(\tilde{V}_0(z_t,n_t;w,r)\) I solve for the policy \(n_{t+1} = \kappa_0(z_t,n_t)\) that maximizes the RHS of the functional equation in 4 taking as given my guess for \(\tilde{V}(\cdot)\) and prices \(w\) and \(r\):

\[
\kappa_0(z_t,n_t) = \arg \max_{n_{t+1}} \pi(z_t,n_t,n_{t+1};w) + \frac{1}{1+r} \tilde{E}[\tilde{V}_0(z_{t+1},n_{t+1};w,r)].
\]

Again, note here that the \(\tilde{E}[\cdot]\) operator takes expectations with respect to the managers’ subjective distribution for future idiosyncratic shocks.

Then I implement Howard’s improvement algorithm by first applying the Bellman operator that imple-
ments the policy $\kappa_0(\cdot)$ for a fixed number $T$ of periods. So for $\tau = 1, 2, ..., T$:

$$
\hat{V}_\tau(z_t, n_t; w, r) = \pi(z_t, n_t, \kappa(z_t, n_t); w, r) + \frac{1}{1+r} \tilde{E}[V_{\tau-1}(z_{t+1}, n_{t+1}; w, r)]
$$

(12)

and finally find a new guess for the optimal policy function $\kappa_1(\cdot)$ by applying the maximization in 11 using guess $\hat{V}_\tau(\cdot)$ as the continuation value. Then if the distance between the previous and current guesses of the policy function is under some pre-specified tolerance I have found the firm’s optimal policy, that is if:

$$
\max_{z, n} \| \kappa_1(z, n) - \kappa_0(z, n) \| < \varepsilon.
$$

In practice I pick $\varepsilon = 10^{-20}$ and $T = 300$. Then it is straightforward to iterate again on the solution to the optimal policy to obtain the ultimate guess for the managers’ subjective valuation of the firm $\hat{V}(\cdot)$ by applying the procedure in equation 12.

If the maximum distance exceeds $\varepsilon$ I instead treat $\kappa_1(\cdot)$ as a new guess of the policy function and re-apply Howard’s improvement algorithm in 12 to obtain a new guess for the policy function.

C.1.3 Computing the stationary distribution of firms across the state space

After solving for managers’ policy function $\kappa(z, n)$ I compute the stationary distribution of firms across the discretized state space $\phi(z, n)$ numerically. Specifically, I exploit the Markovian structure of the model and employ non-stochastic simulation based on Young (2010).

To start, I make a guess that the stationary distribution is uniform across the discrete grid of states $(z, n)$, calling this initial guess $\phi_0(z, n)$. Then, I obtain a new guess $\phi_1(z, n)$ by moving the mass at each point in the state space forward in time according the dynamics of the model. This involves distributing a fraction of the mass currently at point $(z_j, n_k)$ to point $(z_{p}, \kappa(z_j, n_k))$ according to the objective transition probability $Pr(z_{t+1} = z_{p} | z_t = z_n)$. Note that this procedure acknowledges that labor $n$ moves endogenously under managers’ potentially-biased policy function but productivity moves according to its true stochastic process from equation 1.

After moving all of the mass forward I have obtained $\phi_1(z, n)$. I then compute the maximum distance between this new guess and the previous one,

$$
d = \max_{z, n} \| \phi_1(z, n) - \phi_0(z, n) \|
$$

. If $d$ is under a pre-specified tolerance I deem $\phi_1(z, k)$ to be the stationary distribution $\phi(z, k)$. Otherwise, I repeat the procedure iteratively until the distance between $\phi_{\tau+1}(z, k)$ and $\phi_{\tau}(z, k)$ is less than the tolerance.

C.1.4 Computing labor market equilibrium

Given any guess for the wage $w$ I can compute managers’ optimal policies $\kappa(z, n)$ and the stationary distribution of firms $\phi(z, n)$. Using the stationary distribution $\phi(\cdot; w)$ I can then test whether the labor market is in equilibrium. First, I compute the household’s consumption $C = wN^D + \Pi$, where $N^D = \int_{z \times N} n \cdot \phi(z, n; w)dzdn$ is aggregate labor demand and $\Pi$ is the household’s total capital income (see equation 6) under the current guess for the manager’s policies and the wage. Then I find the household’s desired labor supply $N^S$ given $C$ and $w$ according to its intratemporal labor-leisure tradeoff in equation 8. I thus obtain excess labor demand $N^D - N^S$, which equals zero in equilibrium.

To compute the equilibrium of the economy I therefore write a function that computes the economy’s excess labor demand given some wage, namely by solving for managers’ optimization problem, for the stationary distribution and then labor demand and supply given the guess for the wage. I employ a standard nonlinear one-dimensional solver in Matlab to find a wage $w$ for which excess labor demand is close to zero.
C.1.5 Computing moments for the population of firms in the economy

Given I compute the equilibrium stationary distribution of firms $\phi(z, n)$ numerically it is straightforward to use this object to compute population moments for the firms in the model. This procedure avoids drawing random numbers and thus introducing any simulation error.

For illustration, consider any outcome $X(z, n)$ that is a function of the state space in the model. The mean value of $X(z, n)$ is then $E[X(z, n)] = \sum_{z,n} X(z, n) \cdot \phi(z, n)$ where $E[\cdot]$ takes the expectation with respect to the stationary distribution of firms in the model’s equilibrium. For moments of dynamic variables, like the firms’ sales growth that depend say on a firm’s shock next period $z_{t+1}$ in addition to the current state $(z_t, n_t)$ I use the dynamic distribution $\hat{\phi}(z, n, z’) = \phi(z, n) \cdot Pr(z_{t+1} = z’|z_t = z)$ and again computes moments using these weights to average across potential values of the random variable.

C.2 Computing Managers’ Beliefs About Sales Growth Between Quarters $t$ and $t+4$

Relative to the typical dynamic model of firms in heterogeneous-agent macro and corporate finance, my baseline model from Section 3 is actually pretty simple and computationally tractable to solve. Given some parameters, solving for equilibrium typically takes about 10 to 15 seconds on my quad-core 3.6 GHz 2017 iMac with 32GB of RAM.

What is more computationally intensive is obtaining managers’ beliefs about sales growth between quarter $t$ and $t+4$, which are necessary for obtaining moments in the model about managers’ forecast errors that I use to discipline my estimates of managers’ subjective stochastic process, namely $\{\mu, \sigma, \tilde{p}\}$. Specifically, sales in period $t+4$ are a function of the firm’s idiosyncratic shock and the firm’s labor force in period $t+4$:

$$\hat{y}_{t+4} = z_{t+4} n_{t+4}.$$

From the standpoint of quarter $t$ and the firm’s current state $(z_t, n_t)$, this object is a random variable depending on all possible paths of shocks between $t$ and $t+4$, $\zeta = \{z_{t+1}, z_{t+2}, z_{t+3}, z_{t+4}\}$, particularly because

$$n_{t+4} = \kappa(z_{t+3}, n_{t+3}) = \kappa(z_{t+3}, \kappa(z_{t+2}, n_{t+2})) = \kappa(z_{t+3}, \kappa(z_{t+2}, \kappa(z_{t+1}, n_{t+1}))) = \kappa(z_{t+3}, \kappa(z_{t+2}, \kappa(z_{t+1}, \kappa(z_t, n_t))))).$$

So conditional on $(z_t, n_t)$ $\hat{y}_{t+4}$ is a four-dimensional object that occurs with a given probability depending on a possible path of shocks $Pr(\zeta|z_t)$. To compute a managers’ forecast for future sales growth as well as an unbiased forecast thus involves computing two separate expectations with respect to two distinct four-dimensional discrete probability mass functions (depending on whether we forecast with the objective or subjective stochastic process for shocks to $z$). Since my grid for potential $z$-shock values has 21 points, computing a forecast for sales growth between $t$ and $t+4$ for a given point in the state space $(z_t, n_t)$ involves a summation with $21^4 = 194,481$ terms.

In order to keep memory requirements tractable, I thus apply the law iterated expectations liberally to compute forecast error moments in the model. To begin, I store the conditional subjective and objective expectation for sales growth between $t$ and $t+4$

$$E[\Delta y_{t+4} | z_t, n_t] = \sum_{\zeta} Pr(\zeta|z_t) \Delta y_{t+4}(z_t, n_t, \zeta)$$

$$E[\Delta y_{t+4} | z_t, n_t] = \sum_{\zeta} Pr(\zeta|z_t) \Delta y_{t+4}(z_t, n_t, \zeta)$$

$$E[\Delta y_{t+4} | z_t, n_t] = \sum_{\zeta} Pr(\zeta|z_t) \Delta y_{t+4}(z_t, n_t, \zeta)$$

$$E[\Delta y_{t+4} | z_t, n_t] = \sum_{\zeta} Pr(\zeta|z_t) \Delta y_{t+4}(z_t, n_t, \zeta)$$

36
where $\hat{Pr}(\zeta|z)$ and $Pr(\zeta|z)$ denote the subjective and objective probability measures with respect to shocks that might occur between $t$ and $t+4$ conditional on $z_t = z$.

Then I obtain forecast error moments in the population of firms by averaging across firm’s stationary distribution $\phi(z, n)$. For example:

$$E[ForecastError_{t,t+4}] = E \left[ \hat{E}[\Delta y_{t,t+4} - \Delta y_{t,t+4}] \right]$$

$$= E \left[ \hat{E}[\Delta y_{t,t+4}(z, n)] - E[\Delta y_{t,t+4}(z, n)] \right]$$

$$= \sum_{z, n} \phi(z, n) \cdot \left[ \hat{E}[\Delta y_{t,t+4}(z, n)] - E[\Delta y_{t,t+4}(z, n)] \right]$$

where the $E[\cdot]$ operator takes expectations across the state space after conditioning on $(z, n)$ rather than across future shocks. Recall that this moment is particularly useful for pinning down the extent of managerial uncertainty about shocks to $\log(z_t)$, namely $\hat{\mu} - \mu$.

I use a similar procedure to obtain managers’ the subjective mean absolute deviations from their forecast:

$$\text{MAD}(z_t, n_t) = \hat{E} \left[ \|\Delta y_{t,t+4}(z_t, n_t) - \hat{E}[\Delta y_{t,t+4}(z_t, n_t)]\| \right]$$

$$\sum_{\zeta} \hat{Pr}(\zeta|z_t) \cdot \left[ \|\Delta y_{t,t+4}(z_t, n_t, \zeta) - \hat{E}[\Delta y_{t,t+4}(z_t, n_t)]\| \right]$$

and the objective absolute forecast error conditional on $(z_t, n_t)$:

$$E[AbsForecastError_{t,t+4}(z_t, n_t)] = E \left[ \|\Delta y_{t,t+4}(z_t, n_t) - \hat{E}[\Delta y_{t,t+4}]\| \right]$$

$$= \sum Pr(\zeta|z_t) \cdot \left[ \|\Delta y_{t,t+4}(z_t, n_t, \zeta) - \hat{E}[\Delta y_{t,t+4}(z_t, n_t)]\| \right]$$

. Then applying the law of iterated expectations again I ultimately use to compute the mean excess absolute forecast error in the model:

$$E[ExcessAbsForecastError_{t,t+4}] = E \left[ E[AbsForecastError_{t,t+4}(z_t, n_t)] - \text{MAD}(z_t, n_t) \right]$$

$$= \sum_{z, n} \phi(z, n) \cdot \left[ E[AbsForecastError_{t,t+4}(z, n)] - \text{MAD}(z, n) \right]$$

Recall that this moment is my crucial target for disciplining the relative magnitude of managers’ subjective uncertainty about shocks to $\log(z_t)$, $\hat{\sigma}$, relative to the true volatility of those shocks $\sigma$.

I also apply the law of iterated expectations in a similar fashion to compute the final forecast error moment I require for estimation and which helps pin down managers’ perception of shock persistence $\hat{\rho}$ relative to the true persistence $\rho$. Namely I compute the covariance between sales growth between $t - 1$ and $t$ and the forecast error for sales growth between $t$ and $t + 4$:

$$\text{Cov}(\Delta y_t, ForecastError_{t,t+4}) = E \left[ \frac{(\Delta y_t(z_t, n_t) - E[\Delta y_t(z_t, n_t)]) \cdot ForecastError_{t,t+4}(z_t, n_t)}{-E[ForecastError_{t,t+4}(z_t, n_t)]} \right] \quad (13)$$

Although seemingly straightforward, computing this moment is slightly more complicated as $\Delta y_t$ is really a function of $(z_{t-1}, n_{t-1}, z_t)$ so I need to take the expectation $E[\cdot]$ using the distribution $\hat{\phi}(z, n, z') = \phi(z, n) \cdot Pr(z_{t+1} = z'|z_t = z)$. Applying the law of iterated expectations here crucially relies on $\Delta y_t$ being deterministic conditional on $(z_{t-1}, n_{t-1}, z_t)$ which greatly simplifies the number of computations required as we can then separately compute the two terms inside the outermost brackets in equation 13.
C.3 Computing True Firm Value

I compute true firm values by solving the functional equation in 5 numerically using standard dynamic programming results (e.g., see Stokey et al., 1989). Note that the following operator is a contraction for any function \( f(\cdot) \) whose domain is the state space of the economy:

\[
T(f(z_t, n_t) = \pi(z_t, n_t, \kappa(z_t, n_t); w) + \mathbb{E}[f(z_{t+1}, \kappa(z_t, n_t))],
\]

where crucially \( \mathbb{E}[\cdot] \) is the expectations operation with respect to the true stochastic process for \( z_t \), as given in 1 in the main text. Thus, starting from a guess \( V^0(\cdot) \) for \( V(\cdot) \) or \( V^c(\cdot) \), I update the guess by letting \( V^1(z_t, n_t) = T(V^0(z_t, n_t)) \) and iterating until the sup norm between \( V^m(\cdot) \) and \( V^{m+1} \) is under a pre-specified tolerance. This is computationally inexpensive (and arguably trivial), yet it crucially helps me compare managers’ subjective valuations of their own firms \( \tilde{V}(\cdot) \) against the true value \( V(\cdot) \) delivered by managers’ policies.

C.4 GMM Estimation Details

C.4.1 SBU variable definitions

Here I define the specific variables from the SBU that I employ in my structural estimation of the model. Note that although the SBU is a monthly survey in which panel members answer questions about sales and employment every other month, for conformity with my quarterly model I collapse my data to quarterly frequency. In particular, for each quarter I pick the last value reported within a given quarter. I measure all growth rates variables by taking the difference across periods and dividing that by the average, following the long tradition in the literature on business dynamics.

The variables I use in my estimation procedure are the following:

- Sales growth between quarters \( t - 1 \) and \( t \): \( \Delta y_t = \frac{y_t - y_{t-1}}{y_t + y_{t-1}} \)

- Sales growth between quarter \( t \) and \( t + 4 \): \( \Delta y_{t+4} = \frac{y_{t+4} - y_{t}}{y_{t+4} + y_{t}} \)

- Net hiring in period \( t \): \( \Delta n_{t+1} = \frac{n_{t+1} - n_t}{n_{t+1} + n_t} \). Here I take the firm’s employment level in its last response in quarter \( t \) to be \( n_{t+1} \). This treatment is consistent with my one-period lag in between hiring and production in the model and captures real-world lags in recruiting, interviewing and training new employees.

- The firm’s forecast error between \( t \) and \( t + 4 \): \( \text{ForecastError}_{t,t+4} = \tilde{E}[\Delta y_{t+4}] - \Delta y_{t+4} \), with forecasts and realizations measured following the description in Appendix A.2

- The firm’s excess absolute forecast error between \( t \) and \( t + 4 \): \( \text{ExcessAbsForecastError}_{t,t+4} = \|\text{ForecastError}_{t,t+4}\| - \text{MAD}[\Delta y_{t+4}] \), that is the difference between the manager’s realized absolute forecast error and her ex-ante subjective mean absolute deviation: where I again compute the latter according to the description in Appendix A.2.

C.4.2 Computing target moments and the weighting matrix

My estimation targets a vector of eight moments \( m(X) \) described in Section 4.2 of the main text. Table C.1 below reproduces their values, their standard errors and also shows how many firm-quarter observations I use to compute each moment. Since the SBU is a relatively small dataset and I need firms to keep responding to the survey for a year in order to record a forecast error observation, I compute each moment using all firm-quarter observations for which I observe the necessary variables. This means my moments are ultimately not based on the same set of observations, which I take into account when I compute the variance-covariance matrix of moments that I use to construct the moment-weighting matrix of my econometric objective function.
Two of my target moments are means, namely the mean forecast error and mean excess absolute forecast error. I compute this as simple arithmetic means. However, the other six moments are covariances. Since variability of sales, employment, and forecast errors in the data may reflect persistent differences across firms and aggregate shocks, I compute my target variance and covariance moments based only on within-firm variation after controlling for aggregate shocks. That is, I first regress each of the variables that go into one of my variance or covariance targets on a full set of firm and date fixed effects and then compute the target moment on the residuals from those regressions.

As I described in Section 4.2 of the main text, I use the GMM optimal weighting matrix in my econometric minimization procedure, which for my purposes consists of the inverse of the firm-clustered variance-covariance matrix of targeted moments \( \Omega = \mathbb{E}[m(X)m(X)'] \). This treatment of heteroskedasticity accounts for within-firm correlation across observations. I estimate this variance-covariance matrix using the influence function approach from Erickson and Whited (2002). Table C.2 shows my estimate \( \hat{\Omega} \). I justify this choice of weighting matrix given the good small sample performance shown for similar simulation-based estimators in Bazdresch et al. (2017).

C.4.3 Minimizing the econometric objective and computing standard errors

My structural estimation procedure aims to find the vector of parameters \( \theta \) that minimizes the weighted distance between model and data moments, as described briefly in the main text:

\[
\min_{\theta} [m(\theta) - m(X)]'W[m(\theta) - m(X)] .
\]

Recall that I set \( W = \hat{\Omega}^{-1} \), the inverse of the covariance matrix of moments. I conduct this minimization using a standard simulated annealing algorithm that uses randomization to find the minimum of the econometric objective.

Following standard results, as sample sizes go to infinity, the vector of estimated parameters \( \hat{\theta} \) is asymptotically normally distributed with variance \( \Sigma \):

\[
\sqrt{n}(\hat{\theta} - \theta) \to N(0, \Sigma)
\]

where

\[
\Sigma = \left[ \frac{\partial m(\theta)}{\partial \theta'} W \frac{\partial m(\theta)}{\partial \theta} \right]^{-1} \frac{\partial m(\theta)}{\partial \theta'} W \Omega W \frac{\partial m(\theta)}{\partial \theta} \left[ \frac{\partial m(\theta)}{\partial \theta'} W \frac{\partial m(\theta)}{\partial \theta} \right]^{-1}.
\]

In practice, I compute an estimate of the asymptotic variance of by plugging in \( \hat{\Omega} \) in place of \( \Omega \) and obtaining numerical derivatives for \( \frac{\partial m(\theta)}{\partial \theta} \) evaluated at the estimated \( \hat{\theta} \). I compute the latter using two-sided derivatives with step size equal to 2 percent of each element \( \hat{\theta} \) in my baseline calculation:

\[
\hat{\Sigma} = \left[ \frac{\partial m(\theta)}{\partial \theta'} W \frac{\partial m(\theta)}{\partial \theta} \right]^{-1} \frac{\partial m(\theta)}{\partial \theta'} W M \hat{\Omega} W \frac{\partial m(\theta)}{\partial \theta} \left[ \frac{\partial m(\theta)}{\partial \theta'} W \frac{\partial m(\theta)}{\partial \theta} \right]^{-1}.
\]

The matrix \( M = (n_1^{-1}, ..., n_8^{-1})' : (n_1^{-1}, ..., n_8^{-1}) \) where \( n_i \) is the number of observations I use to compute moment \( i = 1, ..., 8 \) in the SBU data. The square root of the diagonal of \( \hat{\Sigma} \) contains the standard errors of the elements in \( \hat{\theta} \).

\footnote{For a couple of moments, specifically those relating sales growth in \( t - 1 \) and \( t \) to forecast errors and sales growth between \( t \) and \( t + 4 \), removing variation due to firm and date fixed effects may introduce dynamic panel complications that could result in biased estimates of those covariances. However, none of these affected moments change by economically significant amounts after the residualizing, so I am not too worried of potential biases due to this dynamic panel structure.}
C.4.4 Sensitivity of estimated parameters to moments

Figure C.1 shows the sensitivity of estimated parameters to moments, which I compute based on Andrews, Gentzkow, and Shapiro (2017). We can see that although moments in the right hand and left hand columns are sensitive to similar sets of moments qualitatively, there are quantitatively significant differences in the sensitivity of moments to parameters across columns. For example, the decreasing returns parameter $\alpha$ and the adjustment costs parameter $\lambda$ are both positively sensitive to the variance of net hiring $\text{Var}(\Delta n_{t+1})$ and the variance of sales growth $\text{Var}(\Delta y_t)$, but $\alpha$ is relatively more sensitive to the second and $\lambda$ the first.

One important feature of Figure C.1 is that it clearly shows how both technological parameters and managers’ subjective beliefs are sensitive to forecast error moments as well as moments concerning sales and employment dynamics with no beliefs. So forecast error moments are also helping me identify the technology parameters and vice versa for moments about firms’ sales and employment dynamics helping me identify technological parameters. Going back to the example with the decreasing returns and adjustment costs parameters $\alpha$ and $\lambda$, both are highly sensitive to the "overextrapolation" moment, the covariance between recent sales growth and the forecast error looking four quarters ahead $\text{Cov}(FE_{t+4}, \Delta y_t)$, although this sensitivity is stronger for $\alpha$ than $\lambda$.

C.5 The Role of General Equilibrium Price Effects

Table C.3 explores how changes in the equilibrium wage – the key price in this economy – fit into the headline results about welfare and reallocation from Tables 9 and 10 in the main text. Looking at the top row of Table C.3, it is now clear that an economy without overconfidence ($\tilde{\sigma} = \sigma$ only) has higher wages and total labor input in equilibrium, which increase the representative consumer’s labor income and – especially – consumption. We see the opposite general equilibrium effects looking at the economy in the second row, which has overconfident managers that don’t overextrapolate ($\tilde{\rho} = \rho$ only). Consumer welfare gains are large here despite lower wages, a modest increase in consumption, and higher profits. In the bottom two rows we that equilibrium wages and the household’s labor and consumption choices differ significantly despite the two economies looking similar in terms of reallocation and gains in consumer welfare in Table 10.

Another interesting result from Table C.3 is that total profits, and therefore managers’ total consumption, declines in three of the four counterfactuals, including the case in which managers are fully unbiased ($\tilde{\rho} = \rho$, $\tilde{\sigma} = \sigma$, and $\tilde{\mu} = \mu$). In all cases the change in profits goes hand in hand with a change in the wage and therefore firms’ operating costs. This result might seem unsettling to the extent that the removing biases does not necessary result in a Pareto improvement. That said, the 4.4 percent reduction in total profits amounts to a reduction in total managerial consumption equal to 0.11 percent of GDP. By comparison, the extra 0.99 consumption-equivalent the household accrues amounts to about 0.96 percent of GDP, almost nine times as much.

The final column of Table C.3 shows the difference in aggregate output (= total revenue less spending on adjustment costs) relative to the baseline biased economy in each of the four counterfactual economies. Although output increases in all cases because my counterfactual economies all reduce the amount of unnecessary spending on adjustment costs, the relative ranking intuitively corresponds to changes in the total amount of labor employed, displayed in the second column. The top row ($\tilde{\sigma} = \sigma$) sees the second largest increase in output despite having the smallest increase in welfare. These results justify my focus on the impact of biases for consumer welfare rather than GDP. Note this is in contrast with much of the misallocation literature, which considers models in which GDP of TFP are equal to welfare (e.g. Hsieh and Klenow, 2009, explicitly say TFP and welfare are one and the same in their model).

C.6 Managerial Biases Interact with Other Public Policies

In this section, I show that the cost of managerial overconfidence and overextrapolation is higher when consumers and firms in my model economy are subject to distortionary payroll and labor income taxes. Similarly, I show that the welfare costs of distortionary taxation are higher in an economy with biased managers, relative to an economy in which managers have rational expectations.

I modify the representative consumer’s budget constraint by introducing labor income taxes $\tau_n$ and rebating the tax revenue using a lump sum transfer $T_t$, while also taxing firms’ wage bill by $\tau_p$. The
following equations show firm’s cash flow, the household’s budget constraint, and the government’s budget balance condition in this setup:

\[
\begin{align*}
\pi(z_t, n_t, n_{t+1}; w_t) &= z_t n_t^\alpha - (1 + \tau_p) w_t n_t - AC(n_t, n_{t+1}) \\
C_t + B_{t+1} &= (1 - \tau_n) w_t N_t + (1 + r_t) B_t + \Pi_t + T_t \\
T_t &= (\tau_n + \tau_p) w_t N_t.
\end{align*}
\]

In Figure C.2 I show how the results from my main macro counterfactual experiment depend on the combination of the labor income and payroll taxes in place in the economy. Namely, each point in the figure compares consumer welfare in an economy with rational managers and taxes relative to an economy in which managers are overconfident and overextrapolate to the extent that I estimate in Section 4.2 of the main text. For each point in the figure, I re-calibrate the household’s disutility of labor \(\chi\) targeting a steady-state quantity of labor \(N\) equal to 1/3 in the equilibrium with biased managers and taxes. The broad lesson from this exercise is that larger distortionary taxes of either kind increase the cost of having biased managers.

Figure C.3 shows the results from a related exercise, looking at how the welfare costs of distortionary income taxes depend on whether managers are biased. Each point in the figure takes an economy with labor income taxes \(\tau_n\) according to its position along the horizontal axis and plots on the vertical axis how much higher consumer welfare would be in the stationary equilibrium with no taxes, \(\tau_n = 0\). As with Figure C.2, I re-calibrate the household’s disutility of labor \(\chi\) targeting \(N = 1/3\) in the "initial" equilibrium with taxes. The two lines in the figure correspond to the welfare costs of the distortionary tax if the economy has biased managers versus not, holding the rest of the parameters fixed at their estimated values. With biased managers, taxes are more costly in terms of consumer welfare.

The intuition for why taxes amplify the cost of managerial biases in Figure C.2 and why managerial biases amplify the costs of distortionary taxes in Figure C.3 is related to the envelope theorem. Namely, when the representative consumer’s consumption and leisure are close to their (undistorted) optimal levels, changing other parameters of her utility maximization problem has second order welfare effects that are relatively small. When consumption and leisure are distorted, the second order effects from changing further distortions like taxes or managerial biases become larger. I view these results as further motivation for why policy-makers should care about pervasive sources of inefficiency like managerial overconfidence and overextrapolation, even if it may be difficult to design policies that change the nature of managerial biases themselves.

### C.7 Investment Model Solution and Estimation

#### C.7.1 Solution

I solve the model of investment dynamics under possibly incorrect subjective beliefs using the same techniques I outline in Section 4 and earlier portions of Appendix C for the baseline specification. Solving this version of the model with static labor choices is actually simpler because the equilibrium wage does not enter into managers’ dynamic investment decision. So I impose the normalization that the function scaling firm revenues \(A(w) = 1\) under the equilibrium wage \(w\). This means I do not need to solve for the economy’s equilibrium wage to compute the stationary distribution of firms across the state space, \(\phi(z, k)\) and hence to compute any moments from that distribution I might need for estimation.

#### C.7.2 Estimation

As with the baseline model, I pick a set of parameters – mainly concerning the household– and estimate the main parameters that govern hiring and firing decisions by minimizing the weighted distance between a set of model and data moments. Table C.4 shows the externally-calibrated parameters for the labor-based model specification. I estimate seven parameters, namely the earnings elasticity of capital, \(\alpha\), the magnitude of partial irreversibility \(\lambda_i\) and quadratic adjustment costs \(\lambda_q\), the subjective and objective persistences of the firm-level shock process \(\rho\) and \(\tilde{\rho}\), as well as the subjective and objective standard deviation of shock
innovations, $\sigma$ and $\tilde{\sigma}$. Finally, I estimate the subjective mean of firm level shocks, $\tilde{\mu}$ (having normalized the true mean $\mu$ to zero).

I estimate the model of investment dynamics by matching the three forecast error moments from the SBU as well as moments that capture the dynamics of investment and output in Compustat data. I use data from two different sources because I do not have reliable data on capital expenditures and capital stocks in the SBU. To mitigate concerns that my two data sources are not conformable, I restrict attention in this baseline estimation to a sample of Compustat firms with less than 7500 employees, the 99th percentile for employment in the SBU. However, I acknowledge that this may not fully address concerns that a sample of public firms may not be representative of my SBU sample or of the US economy (see for example Davis et al., 2007). Since my estimation targets two disjoint sets of moments I choose as weighting matrix the inverse covariance matrix of the moments (assuming zero covariance across datasets), adjusted to place equal weight on Compustat and SBU moments. Adjusting the weight matrix is important in this case because the SBU sample is much smaller than my Compustat sample, so SBU moments would receive little weight in the econometric minimization procedure if I did not adjust.

Tables C.5a and C.5b show my targeted data moments, their model counterparts, and my parameter estimates. Quantitatively, the parameters of the subjective and objective stochastic process are respectively similar to the results from the baseline estimates of Section 4.2 for the baseline labor dynamics model. Managers underestimate the mean shock innovation by $\tilde{\mu} = 0.0005$ (relative to $\mu = 0$), which amounts to about 2 basis points of the true standard deviation of those innovations. However they are severely overconfident, with $\tilde{\sigma} = 0.110$, a bit more than half the true standard deviation of innovations, $\sigma = 0.202$. Additionally, they believe the persistence of shocks to be $\tilde{\rho} = 0.977$, significantly larger than $\rho = 0.859$. This discrepancy leads managers in the model to believe shocks have a half life of about 30 quarters when the true half life is only about 4.5. Thus, managers appear to overextrapolate more severely in this exercise than they do under the baseline results from Section 4.2. I obtain moments about the joint within-firm dynamics of output and capital from Compustat Quarterly covering all years between 1990 and 2017. I choose this long sample in order to minimize issues related to estimating moments on a short panel. Since the firms in my SBU sample are older (about half hired their first paid employee prior to 1970) this longer sample does not seem unreasonable.

I restrict attention to firms incorporated and headquartered in the United States, exclude financials and utilities (SIC 4900 & 6000-6999) as is standard in studies of investment, and drop firms with negative sales, net property, plant, and equipment, short term or long-term debt. Additionally, I drop all firm-quarters with employment over 7500 – the 99th percentile of employment in my SBU sample – to focus on firms that are similar in size to the firms in the SBU. Because my model is tailored toward capturing investment dynamics at firms making day-to-day decisions, I also exclude firms whose sales or net property, plant, and equipment change by a factor of three or more either upward or downward. I acknowledge that this procedure is a crude and simple way of choosing firms in Compustat that are not too different from my SBU sample, given the constraint that I do not have high-quality investment data in the SBU.

The variables from compustat that I consider are net investment ($= \text{the growth rate of net property, plant, and equipment between quarter } t-1 \text{ and } t$), the log-capital-output ratio ($= \log(\text{sales}/\text{lagged net property, plant, and equipment})$) and sales growth between quarter $t$ and $t + 1$. Following the usual convention, I assume that the amount of capital available for production in quarter $t$ is the end-of-period amount reported in quarter $t-1$. In order to isolate within-firm variation and exclude cross-firm heterogeneity that is not a part of my model, I regress each of these variables on a full set of firm and date fixed effects and then compute my moments on the mean-zero residuals from those regressions. Specifically, I consider the six unique elements of the covariance matrix of net investment, the log-sales-to-capital ratio, and sales growth, as well as the autocorrelation of the log-sales-to-capital ratio\footnote{I compute the autocorrelation of the log-sales-to-capital ratio as the OLS coefficient from a regression of the residualized log-sales-to-capital ratio on its lag. It is well known that this sort of dynamic panel regression may be inconsistent when estimated via OLS in this way, especially for highly correlated series, so I also compute the autocorrelation using the consistent estimator proposed in Han and Phillips (2010). Both estimators give essentially the same autocorrelation of 0.8, so I proceed with the OLS-based moment.}. I compute these moments and their variance-covariance matrix using simple GMM and then use these estimates as an input into my structural estimation procedure.
Table C.1: **Target Moments For Estimation**

<table>
<thead>
<tr>
<th>Moment</th>
<th>Value</th>
<th>Standard Error</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean(Forecast Error)</td>
<td>-0.013</td>
<td>0.008</td>
<td>1,256</td>
</tr>
<tr>
<td>Mean(Excess Absolute Forecast Error)</td>
<td>0.143</td>
<td>0.007</td>
<td>1,256</td>
</tr>
<tr>
<td>Cov(Forecast Error_{t,t+4}, Sales Growth_{t-1,t})</td>
<td>0.011</td>
<td>0.003</td>
<td>680</td>
</tr>
<tr>
<td>Var(Sales Growth)</td>
<td>0.060</td>
<td>0.004</td>
<td>2,195</td>
</tr>
<tr>
<td>Var(Net Hiring)</td>
<td>0.019</td>
<td>0.003</td>
<td>2,313</td>
</tr>
<tr>
<td>Cov(Net Hiring, Sales Growth)</td>
<td>0.003</td>
<td>0.001</td>
<td>1,858</td>
</tr>
<tr>
<td>Cov(Sales Growth_{t,t+4}, Sales Growth_{t-1,t})</td>
<td>-0.012</td>
<td>0.003</td>
<td>691</td>
</tr>
<tr>
<td>Cov(Sales Growth_{t,t+4}, Net Hiring_{t,t+1})</td>
<td>-0.002</td>
<td>0.001</td>
<td>686</td>
</tr>
</tbody>
</table>

**Notes:** This table shows the values of the eight target moments I use in my baseline estimation of the model in Sections 3 and 4.2 of the main text, reproduced from Table 6a. Here I additionally report the standard errors of each of the target moments and the number of firm-quarter observations from the SBU I use to compute each moment.

Figure C.1: **Sensitivity of Estimated Parameters to Moments**

**Notes:** This figure shows Andrews-Gentzkow-Shapiro (2017) sensitivities for each of the parameters in the baseline model with respect to targeted moments. Each bar corresponds to the coefficient from a theoretical local regression of parameters on moments, with units expressed in terms of standard deviations.
Table C.2: Variance-Covariance Matrix of Moments

<table>
<thead>
<tr>
<th>Moment</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Mean(Forecast Error)</td>
<td>6.16E-05</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2) Mean(Excess Absolute Forecast Error)</td>
<td>-5.78E-06</td>
<td>4.45E-05</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(3) Cov(Forecast Error&lt;sub&gt;t,t+4&lt;/sub&gt;, Sales Growth&lt;sub&gt;t-1,t&lt;/sub&gt;)</td>
<td>1.81E-06</td>
<td>3.93E-06</td>
<td>6.28E-06</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(4) Var(Sales Growth)</td>
<td>6.67E-07</td>
<td>1.11E-05</td>
<td>1.28E-06</td>
<td>1.89E-05</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(5) Var(Net Hiring)</td>
<td>-6.06E-07</td>
<td>1.23E-06</td>
<td>1.75E-07</td>
<td>2.84E-06</td>
<td>6.01E-06</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(6) Cov(Net Hiring, Sales Growth)</td>
<td>6.53E-07</td>
<td>-3.98E-07</td>
<td>-5.80E-08</td>
<td>9.25E-07</td>
<td>1.04E-06</td>
<td>1.37E-06</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(7) Cov(Sales Growth&lt;sub&gt;t,t+4&lt;/sub&gt;, Sales Growth&lt;sub&gt;t-1,t&lt;/sub&gt;)</td>
<td>-1.73E-06</td>
<td>-4.52E-06</td>
<td>-5.81E-06</td>
<td>-1.86E-06</td>
<td>-7.98E-08</td>
<td>3.19E-08</td>
<td>7.12E-06</td>
<td></td>
</tr>
<tr>
<td>(8) Cov(Sales Growth&lt;sub&gt;t,t+4&lt;/sub&gt;, Net Hiring&lt;sub&gt;t,t+1&lt;/sub&gt;)</td>
<td>2.32E-07</td>
<td>8.47E-07</td>
<td>-6.25E-08</td>
<td>-3.14E-07</td>
<td>-6.37E-07</td>
<td>-3.84E-07</td>
<td>1.51E-07</td>
<td>1.36E-06</td>
</tr>
</tbody>
</table>

Notes: This table shows my estimate of the variance-covariance of the vector of moments targeted in estimation, namely those in Table C.1. I estimate this variance-covariance matrix using the influence function approach from Erickson and Whited (2002).
Table C.3: Impact of Individual Biases: GE Price Effects

<table>
<thead>
<tr>
<th>Counterfactual</th>
<th>Δ C. Welfare %</th>
<th>ΔN %</th>
<th>Δw %</th>
<th>ΔΠ %</th>
<th>ΔC %</th>
<th>ΔY %</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tilde{\sigma} = \sigma$ only</td>
<td>0.40</td>
<td>1.1</td>
<td>4.1</td>
<td>-3.7</td>
<td>0.9</td>
<td>0.8</td>
</tr>
<tr>
<td>$\tilde{\rho} = \rho$ only</td>
<td>0.68</td>
<td>-0.7</td>
<td>-0.6</td>
<td>2.0</td>
<td>0.3</td>
<td>0.4</td>
</tr>
<tr>
<td>$\tilde{\rho} = \rho$ and $\tilde{\sigma} = \sigma$</td>
<td>0.91</td>
<td>0.4</td>
<td>4.7</td>
<td>-3.0</td>
<td>1.1</td>
<td>1.0</td>
</tr>
<tr>
<td>$\tilde{\rho} = \rho$, $\tilde{\sigma} = \sigma$, and $\tilde{\mu} = \mu$</td>
<td>0.99</td>
<td>1.4</td>
<td>6.1</td>
<td>-4.4</td>
<td>1.7</td>
<td>1.6</td>
</tr>
</tbody>
</table>

Notes: This table shows the difference in aggregate consumer welfare, labor, wages, total firm profits, consumption and GDP between an economy whose managers lack one or more of overconfidence ($\tilde{\sigma} = \sigma$), overextrapolation ($\tilde{\rho} = \rho$), or pessimism ($\tilde{\mu} = \mu$) and my baseline economy with biased managers. Each of the economies is at its stationary general equilibrium.

Figure C.2: Taxes Amplify Welfare Impact of Managerial Biases

Notes: This figure shows the welfare change of moving to an economy with rational managers as a function of the payroll and labor income taxes of the baseline economy. For each point in the figure, I re-calibrate the household’s disutility of labor so as to attain aggregate labor $N = 1/3$ in the baseline equilibrium with the combination of taxes in the figure.
Figure C.3: Managerial Biases Amplify Welfare Impact of Taxes

Notes: This figure shows the welfare change of removing labor income taxes, starting from an economy with tax $\tau_n$ and no payroll taxes ($\tau_p = 0$). Each line shows this welfare change depending on whether managers are biased or have rational expectations.
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Description</th>
<th>Target/Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta$</td>
<td>0.026</td>
<td>Quarterly depreciation</td>
<td>NIPA 10% annual</td>
</tr>
<tr>
<td>$\mu$</td>
<td>0</td>
<td>Mean log($z$)</td>
<td>Normalization</td>
</tr>
<tr>
<td>$\hat{\alpha}$</td>
<td>0.275</td>
<td>Revenue elasticity of capital</td>
<td>2/3 labor share in physical output</td>
</tr>
<tr>
<td>$\hat{\nu}$</td>
<td>0.551</td>
<td>Revenue elasticity of labor</td>
<td>2/3 labor share in physical output</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>2</td>
<td>Inverse EIS</td>
<td>Hall (2009)</td>
</tr>
<tr>
<td>$\eta$</td>
<td>2</td>
<td>Inverse Frisch Elasticity of Lab. Supply</td>
<td>Chetty et al. (2011)</td>
</tr>
<tr>
<td>$\beta$</td>
<td>0.96^{1/4}</td>
<td>Household Discount Factor</td>
<td>Annual Interest Rate of 4%</td>
</tr>
<tr>
<td>$\chi$</td>
<td>0.114</td>
<td>Disutility of work</td>
<td>Steady-state Labor $N^{*} = 1/3$</td>
</tr>
<tr>
<td>$\theta$</td>
<td>0.05</td>
<td>Managers’ share of equity</td>
<td>Nikolov and Whited (2014)</td>
</tr>
</tbody>
</table>

**Note:** I set the values for $\hat{\alpha}$ and $\hat{\nu}$, the coefficients on the revenue production function assuming a coefficient of 2/3 in physical output and returns to scale of 0.81 as implied by my estimate of the returns to scale for revenue $\alpha$ in Table C.5b.
Table C.5: **Structural Estimation of Investment Model**

(a) **Data and Model Moments: Investment Model**

<table>
<thead>
<tr>
<th>Moment</th>
<th>Data</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean(Forecast Error)</td>
<td>-0.014</td>
<td>-0.014</td>
</tr>
<tr>
<td>Mean(Excess Absolute Forecast Error)</td>
<td>0.144</td>
<td>0.144</td>
</tr>
<tr>
<td>Cov(Forecast Error$<em>{t,t+4}$, Sales Growth$</em>{t-1,t}$)</td>
<td>0.011</td>
<td>0.012</td>
</tr>
<tr>
<td>Var(Net Investment)</td>
<td>0.020</td>
<td>0.010</td>
</tr>
<tr>
<td>Var(Sales Growth)</td>
<td>0.055</td>
<td>0.044</td>
</tr>
<tr>
<td>Cov(Sales Growth, Net Investment)</td>
<td>0.003</td>
<td>0.002</td>
</tr>
<tr>
<td>Cov(Sales Growth$<em>{t,t+4}$, Sales Growth$</em>{t-1,t}$)</td>
<td>-0.010</td>
<td>0.006</td>
</tr>
<tr>
<td>Cov(Sales Growth$<em>{t,t+4}$, Net Investment$</em>{t}$)</td>
<td>0.004</td>
<td>0.002</td>
</tr>
<tr>
<td>Autocorr(log(Sales/Lagged Capital))</td>
<td>0.803</td>
<td>0.733</td>
</tr>
</tbody>
</table>

(b) **Estimated Parameters: Investment Model**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Explanation</th>
<th>Estimate (SE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>Earnings curvature</td>
<td>0.596 (0.003)</td>
</tr>
<tr>
<td>$\lambda_q$</td>
<td>Quadratic adjustment cost</td>
<td>0.088 (0.003)</td>
</tr>
<tr>
<td>$\lambda_i$</td>
<td>$k$-resale loss</td>
<td>0.091 (0.001)</td>
</tr>
<tr>
<td>$\rho$</td>
<td>True shock persistence</td>
<td>0.864 (0.0009)</td>
</tr>
<tr>
<td>$\tilde{\rho}$</td>
<td>Subjective shock persistence</td>
<td>0.960 (0.0009)</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>True shock volatility</td>
<td>0.200 (0.0002)</td>
</tr>
<tr>
<td>$\tilde{\sigma}$</td>
<td>Subjective shock volatility</td>
<td>0.100 (0.0002)</td>
</tr>
<tr>
<td>$\tilde{\mu}$</td>
<td>Subjective shock mean</td>
<td>-0.001 (0.000007)</td>
</tr>
</tbody>
</table>

**Notes:** This table shows the results from my estimation of the investment model from Appendix B.5, with the estimation described in Appendix C.7.2. Sub-table C.5a (top) shows my target moments in the data and the corresponding model moments after choosing the vectors of parameters that minimize the weighted distance between model and data moments. I compute the top three forecast error moments from SBU data with (the sample period covering 10/2014 to 6/2018. All of my variance and covariance based moments are computed after purging variation attributable to firm and date fixed effects. The investment and sales moments come from a sample of Compustat firms with less than 7500 employees, from a sample covering all years from 1990 to 2017. I compute model moments numerically using the stationary distribution of firms across the $(z,k)$ state space. Sub-table C.5b (bottom) shows the values and standard errors of the estimated parameters. Note that I normalize the true mean of the stochastic process for $\log(z)$ to $\mu = 0$. The weighting matrix is the firm-level clustered covariance matrix of the moments, adjusted to place equal weight on Compustat and SBU moments. I perform the numerical optimization using simulated annealing.
D Heterogeneity in Managerial Oversight, Misbehavior, and Bias

My paper undoubtedly connects to the large literature on agency conflicts between managers and shareholders, although conflicts are not the focus of my analysis and contributions. Indeed, my analysis takes as given that biased managers operate the firms in my model, abstracting from the micro-foundations of how biased individuals end up as managers. There are certainly models in which biased individuals are selected for managerial roles as a result of agency conflicts of various sorts, so in that sense my results help us understand new mechanisms for how agency conflicts can impact firm-level performance. By abstracting from explicit agency conflicts, I also potentially ignore how the interaction between biased managers and shareholders or board members may ultimately shape the dynamic decisions that I observe in the data.

To address some of these concerns, in this Appendix I ask how my model captures the behavior of firms in which managers may be subject to stronger or weaker oversight, in which managers appear better or worse-behaved, or simply more biased based on alternative proxies. To consider differences in oversight, I re-estimate versions of my model splitting the SBU sample by median employment, arguing that small firms in my data are more likely to be owner operated and thus managers less subject to external oversight. I also estimate versions of my investment-based model from Appendix C.7 on subsamples of Compustat firms with good versus bad governance, namely firms with above or below median management entrenchment based on the index from Bebchuk et al. (2008).44 45 To explore managerial empire-building tendencies, I also re-estimate my Compustat-investment model on subsamples of firms with mergers and acquisitions (i.e. for which AQCQ>0) in the eight calendar quarters prior to the current date, versus those for which AQCQ is zero or missing over the same period. Finally, I consider how my model captures the behavior of managers that have previously been identified as being biased. I compare estimates of my investment model based on subsamples of firms that Malmendier and Tate (2015) identify as having "overconfident" CEOs versus not based on CEO stock option exercise behavior. If a CEO exercises any vested stock options within one year of the expiration and the options were at least 40 percent in the money 12 months prior to expiration, Malmendier and Tate (2015) identify this CEO as "overconfident".

Since the SBU data are confidential and I cannot as of now match them to Compustat, for the exercises that focus on public firms data I target the three beliefs moments from the SBU and investment and output moments taken from the relevant subsample of Compustat. Any differences in the estimated parameters and so forth ultimately come from differences in firm-level behavior across the subsamples. If we conjecture that managerial beliefs are less rational in firms with more severe agency conflicts or more badly-behaved managers, we may also expect differences in estimated parameters across subsamples would be more extreme if we could obtain subsample-specific beliefs moments.

Tables D.1 and D.2 shows the results from my estimation across subsamples of firms. Tables D.1a and D.2a show the data and model moments for each estimation. Looking at the data moments on their own, it appears that subsamples with less oversight, empire-building and "overconfident" managers have more erratic or overreactive investment behavior based on the volatility of investment and its covariance with contemporaneous sales growth.

Looking now at Tables D.1b and D.2b, I argue my estimates of model parameters differ in expected ways across sub-samples. In particular, I find the behavior of firms where there appears to be less oversight is consistent with their managers being more biased in the sense of there being a larger gap between $\tilde{\rho}$

---

44 This E-index takes a value between one and six depending on how many governance provisions associated with entrenched management are in place at a particular firm in a particular year. The governance provisions considered include poison pills, golden parachute arrangements, staggered boards, supermajority requirements for mergers, and provisions that make it difficult for shareholders to amend the firm's charter or by-laws.

45 The index is not time-invariant but about 80 percent of the variation is across firms, so I construct my "good" and "bad" governance subsamples by first averaging the firm's index score across years and then splitting the sample by the median score.

46 Note that this meaning of overconfident is conceptually distinct from the precise meaning in my paper. Malmendier and Tate (2015) use the word "overconfident" to describe CEOs who appear to believe their firm's future performance is inherently better than it appears currently. I use "overconfident" to describe managers who underestimate the volatility of future business conditions, that is managers who feel less uncertain about the firm's future performance than they should be. What Malmendier and Tate (2015) term overconfident is more similar to the concept of "over-optimism" in my work.
and $\rho$ and a smaller $\tilde{\sigma}/\sigma$ ratio. Other parameters also behave in intuitive ways. In the case of firms with weak governance, managers also face somewhat weaker adjustment frictions, especially in the form of partial irreversibility, possibly because they need to devote less managerial time and fewer resources negotiating with the board and shareholders about their plans and strategies. By contrast, firms that recently engaged in M&A activity have similar or perhaps slightly higher adjustment costs. This result could be an indication that managers with empire-building tendencies are especially willing to spend on investment projects and M&A despite those actions being relatively costly. Finally, firms with "overconfident" CEOs according to Malmendier and Tate (2015) also face somewhat smaller adjustment costs and behave in ways consistent with them being significantly more biased than firms with CEOs who do not appear "overconfident".

Finally, looking at Table D.3, I ask whether more badly-behaved managers destroy a larger share of firm value than those who appear to be better behaved or subject to less severe agency conflicts. I find mixed results. Public firms with highly entrenched management could increase firm value by 3.3 percent if they hired a rational manager, less than the 4.1 percent for firms with less entrenched management. This result owes to the lower magnitude of adjustment costs in highly-entrenched firms, and also to smaller shocks experienced by these firms ($\sigma = 0.158$ relative to $\sigma = 0.187$ at firms with and low entrenchment). Managers at firms with high entrenchment and poor governance destroy relatively less value because their overreactions are less costly, and because they have fewer opportunities to overreact since the firm is less volatile to begin with. Looking instead at firms conducting versus not conducting M&A in recent quarters, moving to a rational manager increases firm value by 3.3 percent for acquirors and just a bit less, 2.8 percent for non-acquirors. This result is consistent with firms with a taste for acquisitions behaving in a way consistent with more severe overconfidence and overextrapolation. Finally, for firms with "overconfident" CEOs according to Malmendier and Tate (2015), replacing those CEOs with others who have rational expectations would increase the value of the firm by 3.7 percent, more than the 3.3 percent in the subsample for which CEOs don’t appear "overconfident." This is despite firms with "overconfident" CEOs being significantly less volatile and facing smaller adjustment costs. From this analysis, it is clear that simply having a CEO who is more biased does not automatically imply she destroys more firm value. The magnitude of other parameters that help explain the behavior of this biased and potentially badly-behaved CEO also matter, which depends on the type of bad behavior considered and how any particular model captures the behavior.
Table D.1: Sample Split Estimation Results: Large vs. Small SBU Firms

(a) Model and Data Moments for Small vs. Large SBU Firms

<table>
<thead>
<tr>
<th>Moment</th>
<th>Small Data</th>
<th>Small Model</th>
<th>Large Data</th>
<th>Large Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean(Forecast Error)</td>
<td>-0.020</td>
<td>-0.015</td>
<td>-0.006</td>
<td>-0.005</td>
</tr>
<tr>
<td>Mean(Excess Absolute Forecast Error)</td>
<td>0.165</td>
<td>0.158</td>
<td>0.126</td>
<td>0.114</td>
</tr>
<tr>
<td>Cov(Forecast Error_{t,t+4}, Sales Growth_{t-1,t})</td>
<td>0.015</td>
<td>0.009</td>
<td>0.009</td>
<td>0.004</td>
</tr>
<tr>
<td>Var(Sales Growth)</td>
<td>0.075</td>
<td>0.060</td>
<td>0.051</td>
<td>0.038</td>
</tr>
<tr>
<td>Var(Net Hiring)</td>
<td>0.026</td>
<td>0.001</td>
<td>0.014</td>
<td>0.002</td>
</tr>
<tr>
<td>Cov(Net Hiring, Sales Growth)</td>
<td>0.004</td>
<td>0.004</td>
<td>0.001</td>
<td>0.004</td>
</tr>
<tr>
<td>Cov(Sales Growth_{t,t+4}, Sales Growth_{t-1,t})</td>
<td>-0.017</td>
<td>-0.018</td>
<td>-0.011</td>
<td>-0.005</td>
</tr>
<tr>
<td>Cov(Sales Growth_{t,t+4}, Net Hiring_{t,t+1})</td>
<td>-0.002</td>
<td>-0.004</td>
<td>-0.001</td>
<td>-0.002</td>
</tr>
</tbody>
</table>

(b) Parameter Estimates for Large vs. Small SBU Firms

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Small</th>
<th>Large</th>
</tr>
</thead>
<tbody>
<tr>
<td>α</td>
<td>0.611 (0.089)</td>
<td>0.588 (0.113)</td>
</tr>
<tr>
<td>λ</td>
<td>28.707 (1.423)</td>
<td>24.084 (2.368)</td>
</tr>
<tr>
<td>ρ</td>
<td>0.752 (0.008)</td>
<td>0.864 (0.011)</td>
</tr>
<tr>
<td>ρ̃</td>
<td>0.889 (0.007)</td>
<td>0.924 (0.013)</td>
</tr>
<tr>
<td>σ</td>
<td>0.232 (0.001)</td>
<td>0.190 (0.001)</td>
</tr>
<tr>
<td>σ̃</td>
<td>0.086 (0.002)</td>
<td>0.099 (0.002)</td>
</tr>
<tr>
<td>μ̃</td>
<td>-0.004 (0.0001)</td>
<td>-0.001 (0.0001)</td>
</tr>
</tbody>
</table>

Notes: This table shows the results from my structural estimation of the model from Section 3 splitting the sample by median employment. Sub-table D.1a (top) shows my target moments in the data and the corresponding model moments for each sub-sample after choosing the vector of parameters that minimize the weighted distance between model and data moments. I estimate all data moments using SBU data with the sample period covering 10/2014 to 6/2018. All of the variances and covariances I target correspond to within-firm variation. Namely, before computing my target covariances and variances I regress all observations of a full set of firm and date fixed effects to purge variation due to aggregate shocks and persistent differences across firms and then compute the variances and covariances on the residual of those regressions. I compute model moments numerically from the stationary distribution of firms across the \((z, n)\) state space of the model. Sub-table D.1b (bottom) shows the values and standard errors of the parameters that minimize the weighted distance between model and data moments for each of the subsamples. Note that I normalize the true mean of the stochastic process for \(\log(z)\) to \(\mu = 0\). My choice of weighting matrix is the firm-level clustered covariance matrix of SBU data moments, namely the GMM efficient weighting matrix. I perform the numerical optimization of the econometric objective using a simulated annealing algorithm.
Table D.2: Sample Split Estimation Results: Compustat

(a) Model and Data Moments for Sample-Split Estimations

<table>
<thead>
<tr>
<th>Moment</th>
<th>Entrenchment</th>
<th>Acquisitions</th>
<th>&quot;Overconfident&quot; CEO</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>High</td>
<td>Low</td>
<td>Data</td>
</tr>
<tr>
<td>Mean(Excess Absolute Forecast Error)</td>
<td>0.144</td>
<td>0.144</td>
<td>0.144</td>
</tr>
<tr>
<td>Cov(Forecast Error&lt;sub&gt;t,t+4&lt;/sub&gt;, Sales Growth&lt;sub&gt;t-1,t&lt;/sub&gt;)</td>
<td>0.011</td>
<td>0.011</td>
<td>0.011</td>
</tr>
<tr>
<td>Var(Sales Growth)</td>
<td>0.026</td>
<td>0.028</td>
<td>0.035</td>
</tr>
<tr>
<td>Cov(Sales Growth&lt;sub&gt;t,t+4&lt;/sub&gt;, Net Investment&lt;sub&gt;t&lt;/sub&gt;)</td>
<td>0.002</td>
<td>-0.000</td>
<td>0.002</td>
</tr>
<tr>
<td>Cov(Sales Growth&lt;sub&gt;t,t+4&lt;/sub&gt;, Sales Growth&lt;sub&gt;t−1,t&lt;/sub&gt;)</td>
<td>-0.001</td>
<td>-0.002</td>
<td>-0.001</td>
</tr>
<tr>
<td>Cov(Sales Growth&lt;sub&gt;t,t+4&lt;/sub&gt;, Net Investment&lt;sub&gt;t+4&lt;/sub&gt;, Net Investment&lt;sub&gt;t&lt;/sub&gt;)</td>
<td>-0.000</td>
<td>-0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Autocorr(log(Sales/Lagged Capital))</td>
<td>0.839</td>
<td>0.698</td>
<td>0.859</td>
</tr>
</tbody>
</table>

(b) Parameter Estimates for Compustat Sample-Split Estimations

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Entrenchment</th>
<th>Acquisitions</th>
<th>&quot;Overconfident&quot; CEO</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>High</td>
<td>Low</td>
<td>Yes</td>
</tr>
<tr>
<td>α</td>
<td>0.601 (0.014)</td>
<td>0.591 (0.011)</td>
<td>0.601 (0.049)</td>
</tr>
<tr>
<td>λ&lt;sub&gt;q&lt;/sub&gt;</td>
<td>0.154 (0.010)</td>
<td>0.121 (0.002)</td>
<td>0.089 (0.093)</td>
</tr>
<tr>
<td>λ&lt;sub&gt;i&lt;/sub&gt;</td>
<td>0.103 (0.006)</td>
<td>0.131 (0.003)</td>
<td>0.128 (0.006)</td>
</tr>
<tr>
<td>ρ</td>
<td>0.805 (0.003)</td>
<td>0.863 (0.003)</td>
<td>0.831 (0.012)</td>
</tr>
<tr>
<td>β</td>
<td>0.965 (0.006)</td>
<td>0.969 (0.001)</td>
<td>0.959 (0.008)</td>
</tr>
<tr>
<td>σ</td>
<td>0.158 (0.001)</td>
<td>0.187 (0.0003)</td>
<td>0.182 (0.001)</td>
</tr>
<tr>
<td>γ</td>
<td>0.062 (0.003)</td>
<td>0.089 (0.0006)</td>
<td>0.079 (0.002)</td>
</tr>
<tr>
<td>μ</td>
<td>-0.002 (0.0001)</td>
<td>-0.002 (0.0001)</td>
<td>-0.001 (0.0004)</td>
</tr>
</tbody>
</table>

Notes: Table D.2a (top) shows data and model moments from separate estimations of the capital-based model specification for subsamples of Compustat rms with highly-entrenched vs. not highly-entrenched management Bebchuk et al. (2008), a subsample of firms with positive acquisitions (AQCQ) in the past eight quarters versus not, and "overconfident" or "longholder"CEOs versus not according to Malmendier and Tate (2015). Table D.2b (bottom) shows parameter estimates for the capital-based model specification for the same subsamples of Compustat firms.
Table D.3: Change in Firm Value from Hiring Unbiased Manager: Sample Splits

<table>
<thead>
<tr>
<th>Counterfactual</th>
<th>Δ True Firm Value (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Size</td>
</tr>
<tr>
<td></td>
<td>Small</td>
</tr>
<tr>
<td>( \tilde{\rho} = \rho, \tilde{\sigma} = \sigma, \text{ and } \tilde{\mu} = \mu )</td>
<td>2.0</td>
</tr>
</tbody>
</table>

Notes: This table shows percent change in firm value from replacing a biased manager with an unbiased one based on the model estimates for subsamples of SBU firms that are small versus large (i.e. below versus above median employment in the SBU), as well as subsamples of Compustat with highly-entrenched versus not highly-entrenched management Bebchuk et al. (2008), a subsample of firms with positive acquisitions (AQCC) in the past eight quarters versus not, and "overconfident" or "longholder" CEOs versus not according to Malmendier and Tate (2015).