

CEO Overconfidence and Firm-Level Total Factor Productivity

Abstract

I discuss a model to explain how CEO overconfidence might affect the firm productivity. The model shows that overconfident managers have stronger beliefs about their firms' prospects, inducing them to exert higher efforts. This, in turn, increases the probability that the firms' projects can be successful and consequently boosts the firm productivity. The prediction of the model is validated by empirical evidence. I use a sample of S&P 1,500 companies from 1992 to 2023 and follow the literature to measure the productivity at the firm level and managerial overconfidence. The empirical results support the model prediction. The results are robust across different model specifications, firm-fixed effects and within firm analysis.

Keywords: CEO overconfidence; Firm-level total factor productivity; Career concern model.

1 Introduction

In 1983, Steve Jobs gave a revolutionary speech at the International Design Conference in Aspen. He started the talk by asking how many people own an Apple or any personal computer. Few people raised their hands. At that time, personal computers were new and most American households only had televisions. In the speech, Steve Jobs discussed the revolution of personal computers and imagined a future market for computers in every household.¹ By 2021, 95 percent of U.S. households had at least one computer.² In existing finance literature, managers like Steve Jobs, while clearly visionary, are often characterized as overconfident leaders (Hirshleifer et al. 2012).

On one hand, managerial overconfidence can be considered a negative factor to a firm. Overconfident CEOs may overinvest (even in negative net present value projects), especially when their firms have abundant internal resources, or they may underinvest (by forgoing value-enhancing projects), being reluctant to dilute equity when firms need to obtain external funds (Heaton 2002; Malmendier and Tate 2005). On the other hand, empirical evidence shows that overconfident managers can have positive impacts on firms by being greater innovators (Galasso and Simcoe 2011; Hirshleifer et al. 2012). Overconfident CEOs, having stronger beliefs about the firm's future prospects, are also better able to induce other stakeholders, such as suppliers and employees, to make longer commitments to the firm (Phua et al. 2018).

In this paper, I argue that the positive effects that overconfident CEOs can bring to their firms may improve the efficiency of how their firms use resources. Specifically, I add to the literature by investigating the relation between managerial overconfidence and firm-level total factor productivity (TFP). When first developed, TFP was a macroeconomic concept that was often related to the economic growth of a nation (Griliches 1996; Hulten 2001). More recently, the metric has been used to measure productivity at the industry and firm levels (Olley and Pakes 1996;

¹ "The Objects of Our Life. Steve's talk at the 1983 International Design Conference in Aspen" accessed from <https://stevejobsarchive.com/exhibits/objects-of-our-life> on Wednesday, May 21, 2025.

² United States Census Bureau (2024). Computer and Internet Use in the United States: 2021. Accessed from <https://www.census.gov/newsroom/press-releases/2024/computer-internet-use-2021.html> on Wednesday, May 21, 2025.

Levinsohn and Petrin 2003; Brynjolfsson and Hitt 2003; Javorcik 2004; Akerberg et al. 2006; İmrohoroglu and Tüzel 2014; Parrotta et al. 2014; Nagle 2019). The extant literature has shed light on how managerial overconfidence can affect firm inputs such as labor, materials, and investment (Heaton 2002; Malmendier and Tate 2005; Malmendier and Tate 2008; Gervais et al. 2011; Ferris et al. 2013; Phua et al. 2018) and firm outcomes such as firm value and firm innovation (Goel and Thakor 2008; Campbell et al. 2011; Galasso and Simcoe 2011; Hirshleifer et al. 2012). I expand on these streams of literature by combining them to determine how managerial overconfidence affects firm-level TFP.

I adapt the models of Aghion et al. (2009) and Galasso and Simcoe (2011) to theoretically explain how CEO overconfidence might affect firm-level TFP. In this simple two-period model, the CEO of a firm is considering a new project and has compensation based on the project's cash flow, which takes one of two possible values in the second period, initially unknown to both the market and the CEO. CEO overconfidence is characterized by systematically overestimating the probability of the higher cash flow outcome. The CEO then decides how much effort to exert in period 1 based on their beliefs. At the end of period 1, productivity is realized and the market updates its assessment on the project's cash flow. The model predicts that overconfident CEOs have stronger beliefs about firms' prospects, which inspires them to put forth higher effort, subsequently boosting firm productivity.

I also provide empirical evidence to support the prediction of the model. Following the literature (Levinsohn and Petrin 2003), I employ the Levinsohn and Petrin (LP) estimation method to measure firm productivity. To measure managerial overconfidence, I use variables based on the CEO's stock option-exercising and net-buying behavior proposed by Malmendier and Tate (2005). The empirical results show that CEO overconfidence does have a positive effect on firm productivity. The results are robust across different research strategies, including with and without firm fixed effects, controlling for R&D expenses, dealing with missing R&D values, and applying a within-firm-CEO change analysis.

The remainder of the paper is organized as follows. Section 2 discusses relevant literature. I present the model in Section 3 to connect managerial overconfidence and firm productivity. Section 4 describes data and research methodology, followed by empirical results presented in Section 5. Section 6 concludes the paper.

2 Literature Review

2.1 Total Factor Productivity

2.1.1 The National Income Approach

Total factor productivity is a measure of output per unit input originating from the national income measurement tradition (Copeland, 1937; Copeland and Martin, 1938; Stigler, 1947; Barton and Cooper, 1948; Schmookler, 1952; Fabricant, 1954; Ruttan, 1954, 1956; Kendrick, 1955, 1956; and Abramovitz, 1956).³ The basic idea of the metric is to balance the product market and the factor market. As such, the total value of the goods produced in the product market, $p_t Q_t$, where p_t and Q_t are respectively the price and quantity produced at time t , should be equal to the total payments to the inputs in the factor market, $w_t L_t + r_t K_t$, where L_t and K_t respectively represent the labor and capital inputs while w_t and r_t indicates the labor's wage and the capital's rent respectively (Hulten, 2001). Therefore, the following fundamental national accounting identity is established:

$$p_t Q_t = w_t L_t + r_t K_t \quad (1)$$

Equation 1 simply measures GDP in current prices of an economy, implying that when current prices (i.e., p_t , w_t , r_t) change, both sides of the Equation 1 change accordingly, and therefore, the equation is not appropriate to capture economic progress. Hence, to reasonably measure the productivity of economic activity, Equation 1 should use the prices in a base year to value current output and input (i.e., keeping the prices constant). Based on this intuition, the following revision of Equation 1 is introduced (Hulten, 2001):

$$p_0 Q_t = S_t (w_0 L_t + r_0 K_t) \quad (2)$$

Equation 2 differs from Equation 1 in two important aspects. First, Equation 2 uses the prices (i.e., p_0 , w_0 , r_0) in the base year 0. As such, the left-hand side of Equation 2, $p_0 Q_t$, measures the real output and the term in parentheses on the right-hand side, $w_0 L_t + r_0 K_t$, measures the real inputs. Second, Equation 2 introduces a scaling factor, S_t . The intuition behind the factor S_t is that if the real output's growth exceeds the growth in real inputs, then the discrepancy between the two

³ Please see Griliches (1996) and Hulten (2001) for comprehensive review.

clearly indicates an improvement in productivity, which is captured by S_t . By rearranging Equation 2, this intuition becomes more obvious:

$$S_t = \frac{p_0 Q_t}{w_0 L_t + r_0 K_t} \quad (3)$$

Equation 3 implies that S_t can be regarded as the output per unit input index used to measure the productivity of inputs (Hulten, 2001).

Under the national income approach, the output per unit input index is simple to use, however, as Hulten (2001) summarizes, the index is subject to several issues. The substitution bias which is well known in a Laspeyres index is also applicable to the metric. Substitution bias occurs when there are changes in relative prices and producers (or consumers) replace the more expensive goods with the relatively cheaper ones. Another problem is related to how one can interpret the index. As Hulten (2001) mentions, the interpretation of the index can be understood from perspectives of either the producer or the consumer, which leads to a fundamental ambiguity of the index.

2.1.2 The Production Function Approach

The measure of TFP that is most commonly used in the literature is based on the production function (Griliches, 1996; and Hulten, 2001). Tinbergen (1942) used a Cobb-Douglas function to study “technical development”, which is closely related to concept of TFP. However, Solow (1957) started with the following aggregate production function:

$$Q = A(t)f(K, L) \quad (4)$$

In Equation 4, Q , K , and L respectively measure physical outputs, capital and labor inputs. t is time and $A(t)$ refers to the “technical change” and measures “the cumulated effects of shifts over time” (Solow 1957, p. 312). Taking the derivative with respect to t on both sides of Equation 4 yields:

$$\begin{aligned} \frac{\partial Q}{\partial t} &= \frac{\partial A(t)}{\partial t} f(K, L) + A(t) \frac{\partial f(K, L)}{\partial t} \\ \Leftrightarrow \frac{\partial Q}{\partial t} &= \frac{\partial A(t)}{\partial t} f(K, L) + A(t) \left(\frac{\partial f(K, L)}{\partial K} \frac{\partial K}{\partial t} + \frac{\partial f(K, L)}{\partial L} \frac{\partial L}{\partial t} \right) \end{aligned}$$

$$\Leftrightarrow \frac{\partial Q}{\partial t} = \frac{\partial A(t)}{\partial t} f(K, L) + A(t) \frac{\partial f(K, L)}{\partial K} \frac{\partial K}{\partial t} + A(t) \frac{\partial f(K, L)}{\partial L} \frac{\partial L}{\partial t} \quad (5)$$

In addition, from Equation 4:

$$f(K, L) = \frac{Q}{A(t)} \quad (6)$$

And:

$$\frac{\partial Q}{\partial K} = \frac{\partial A(t)}{\partial K} f(K, L) + A(t) \frac{\partial f(K, L)}{\partial K} = A(t) \frac{\partial f(K, L)}{\partial K} \quad (7)$$

$$\frac{\partial Q}{\partial L} = \frac{\partial A(t)}{\partial L} f(K, L) + A(t) \frac{\partial f(K, L)}{\partial L} = A(t) \frac{\partial f(K, L)}{\partial L} \quad (8)$$

Substituting (6), (7), and (8) into (5) gives:

$$\frac{\partial Q}{\partial t} = \frac{\partial A(t)}{\partial t} \frac{Q}{A(t)} + \frac{\partial Q}{\partial K} \frac{\partial K}{\partial t} + \frac{\partial Q}{\partial L} \frac{\partial L}{\partial t} \quad (9)$$

Denote $\dot{Q} = \frac{\partial Q}{\partial t}$, $\dot{A} = \frac{\partial A(t)}{\partial t}$, $\dot{K} = \frac{\partial K}{\partial t}$, and $\dot{L} = \frac{\partial L}{\partial t}$ then Equation 9 becomes:

$$\dot{Q} = \dot{A} \frac{Q}{A(t)} + \frac{\partial Q}{\partial K} \dot{K} + \frac{\partial Q}{\partial L} \dot{L} \quad (10)$$

Divide both sides of (10) by Q :

$$\frac{\dot{Q}}{Q} = \frac{\dot{A}}{A(t)} + \frac{\partial Q}{\partial K} \frac{K}{Q} \frac{\dot{K}}{K} + \frac{\partial Q}{\partial L} \frac{L}{Q} \frac{\dot{L}}{L} \quad (11)$$

In Equation 11, the two terms, $\frac{\partial Q}{\partial K} \frac{K}{Q}$ and $\frac{\partial Q}{\partial L} \frac{L}{Q}$, respectively measure the output elasticities or the relative shares of capital and labor (Solow 1957; Hulten 2001). Denote $s_K = \frac{\partial Q}{\partial K} \frac{K}{Q}$ and $s_L = \frac{\partial Q}{\partial L} \frac{L}{Q}$ and rearrange Equation 11, one gets:

$$\mathcal{R} = \frac{\dot{A}}{A(t)} = \frac{\dot{Q}}{Q} - s_K \frac{\dot{K}}{K} - s_L \frac{\dot{L}}{L} \quad (12)$$

In Equation 12, $\mathcal{R}$ is the Solow residual and is the theoretical base to measure the TFP (Hulten, 2001). To empirically estimate the TFP, economists often write Equation 4 in terms of a Cobb-

Douglas production function and then use the logarithm transformation to estimate the residual from a regression of the log form of the output (i.e., the dependent variable) on the log form of the inputs (i.e., the independent variables) (Wolf 1996; Olley and Pakes 1996). The estimation can be carried out by simply using the Ordinary Least Square (OLS). However, the OLS estimation of the TFP is subject to endogeneity issues (Van Beveren, 2012). Olley and Pakes (1996) provide an estimation algorithm to overcome the selection and simultaneity biases that often emerge when estimating the productivity by introducing an investment proxy to control for the correlation between input variables and unobserved factors when regressing the outputs on the inputs. Levinsohn and Petrin (2003) extend their work by showing that intermediate inputs such as materials or electricity can also solve these issues.

TFP was initially a macroeconomic concept, but it was later used to measure productivity at the industry and the firm levels (Olley and Pakes 1996; Levinsohn and Petrin 2003; Brynjolfsson and Hitt 2003; Javorcik 2004; İmrohoroglu and Tüzel 2014). Researchers have pointed out how firm characteristics can affect TFP (Javorcik 2004; İmrohoroglu and Tüzel 2014) and how TFP can affect firm outcomes such as firm return (İmrohoroglu and Tüzel 2014). However, the linkages between manager characteristics and the firm-level productivity are still underexplored.

2.2 Managerial Overconfidence and the Firm's Production Process

In the literature, researchers have investigated how managerial overconfidence can affect firms' production processes. In particular, overconfident managers can affect both inputs and outcomes. In the following subsections, I discuss relevant literature to review how managerial overconfidence can affect decisions related to the firm's production function.

2.2.1 Managerial Overconfidence and the Firm's Inputs

Researchers have put a substantial effort into studying how managerial overconfidence can affect a firm's investment decisions. Heaton (2002) constructs a model to argue that optimistic CEOs believe outside investors underestimate the value of their firms, making obtaining external funding appear to be too costly, leading them to give up positive net present value projects. In addition, optimistic CEOs also tend to overestimate the values of their own corporate projects, inducing them to invest even in negative net present value projects, if they have sufficient internal funds. Malmendier and Tate (2005) theoretically and empirically show that when firms have sufficient

internal funds, overconfident managers tend to overinvest. Conversely, when firms require external funds, overconfident CEOs tend to underinvest.

In another paper, Malmendier and Tate (2008) point out that overconfident managers tend to overpay for target companies and often make value-destroying merger decisions. This effect is intensified when firms have abundant internal financing. In accordance with this finding, Ferris et al. (2013) demonstrate that overconfident CEOs are more likely to participate in international acquisitions and more likely to use internal resources to finance merger deals.

Gervais et al. (2011) model how firms design compensation contracts to affect how overconfident managers make capital budgeting decisions. They show that by designing a flatter compensation schedule, firms can induce overconfident managers to follow an optimal risky investment policy and exert effort to improve the success rate and value of the firms' investment projects.

Overconfident managers can affect not only the firm investment but also the supply side of the other inputs such as materials or labor. Phua et al. (2018) argue that overconfident managers are better leaders. These CEOs have stronger beliefs about their firm's prospects. This belief can induce stakeholders to make long-term commitments with the firm. In particular, overconfident managers can negotiate with suppliers to make more relationship-specific investments and increase relationship duration. In addition, employees in firms run by overconfident CEOs also tend to stay longer and hold more of the firm's stock in their benefit plans.

2.2.2 Managerial Overconfidence and Firm Outcomes

In finance, the most important outcome to shareholders might be firm value. Goel and Thakor (2008) build a model to show that there is a nonmonotonic relationship between CEO overconfidence and firm value, implying that CEO overconfidence tends to increase firm value initially, but eventually it decreases firm value as the degree of overconfidence exceeds a certain threshold. Therefore, Goel and Thakor (2008) predict that under a value-maximizing corporate governance model, CEOs who are either overly diffident or overly overconfident will often be fired. Only CEOs with a moderate degree of overconfidence are selected to the benefit of shareholders. Consistent with this view, Campbell et al. (2011) theoretically show that firm value is concave in managerial optimism. That means there is an interior optimum level of managerial optimism such that those managers will choose the first-best investment level that maximizes firm value. Hence, their model implies that CEOs with moderate levels of optimism are more likely to

be recruited whereas relatively low- or high-optimism CEOs are more likely to be forced out. They also provide empirical evidence to confirm the prediction.

Overconfident CEOs can not only increase firm value, but they can also positively affect firm innovation. Galasso and Simcoe (2011) develop a model to show that overconfident managers are more likely to innovate than their less confident counterparts. In addition, when product market competition is more intense, the impact of overconfidence on innovation is enhanced. They confirm the model's prediction using a sample of 450 large publicly traded U.S. companies from 1980 to 1994.

Hirshleifer et al. (2012) also point out that overconfident managers are better innovators. Specifically, overconfident CEOs tend to invest more in innovation, and they generate more patents and patent citations. Overconfident managers also use research and development expenditures more efficiently. However, the authors note that greater innovation in firms with overconfident managers is only found in innovative industries. Furthermore, firms run by overconfident CEOs tend to exhibit greater return volatility, suggesting increased firm-specific risk.

Therefore, the existing literature suggests that overconfident managers impact both firm inputs and outcomes. From this theory there might be a potential connection between CEO overconfidence and firm-level TFP. In the next section, I construct a model to explain how CEO overconfidence can affect productivity at the firm level.

3 Model

Aghion et al. (2009) build a variant of the manager's career concern model proposed by Holmstrom (1982) to explain how a manager decides whether to innovate or not. Galasso and Simcoe (2011) extend the model by adding managerial overconfidence to show how an overconfident manager makes innovative decisions. In this section, I adapt the models of Aghion et al. (2009) and Galasso and Simcoe (2011) to link CEO overconfidence to firm-level TFP. When applicable, I clarify the differences between the model presented here and their models.

Following Heaton (2002), Malmendier and Tate (2005), Goel and Thakor (2008), Aghion et al. (2009), and Galasso and Simcoe (2011), I assume there are two time periods, the first period is from $t = 0$ to $t = 1$, and the second period is from $t = 1$ to $t = 2$. The firm is run by a manager who

is considering a new project. In alignment with Heaton (2002) and Gervais et al. (2011), I assume the project's ending cash flow $\mathbb{C}$ can have two possible outcomes, that is, without loss of generality, $\mathbb{C} \in \{0, \bar{\mathbb{C}}\}$ where $\bar{\mathbb{C}} > 0$. At the beginning of period 1 (i.e., at $t = 0$), $\mathbb{C}$ is unknown to both the CEO and the market. In the market, other companies have similar projects, so if the manager leaves the firm and goes to a new firm, he deals with the same situation. Again, without loss of generality, the market's prior beliefs about the project's end of cash flow are assumed to be:

$$Pr_M(\mathbb{C} = \bar{\mathbb{C}}) = Pr_M(\mathbb{C} = 0) = \frac{1}{2}$$

The CEO's prior beliefs about $\mathbb{C}$ may differ from those of the market. If a CEO is overconfident, he is assumed to be optimistic about the good outcome of the project (Heaton 2002; Malmendier and Tate 2005), and his beliefs about $\mathbb{C}$ are characterized as follows:

$$Pr_O(\mathbb{C} = \bar{\mathbb{C}}) = \frac{1}{2}(1 + \Delta)$$

$$Pr_O(\mathbb{C} = 0) = 1 - \frac{1}{2}(1 + \Delta) = \frac{1}{2}(1 - \Delta)$$

where $0 < \Delta < 1$ is a factor to represent CEO overconfidence. In their model, Galasso and Simcoe (2011) define an overconfident CEO as the one who overestimates his ability. In this model, I follow Heaton (2002) and Malmendier and Tate (2005) and identify an overconfident manager as one who assesses a higher likelihood of the project's good outcome.

During period 1 (i.e., from $t = 0$ to $t = 1$), the manager decides how much effort to exert.⁴ The effort expended by the CEO is denoted by e . Putting forth effort leads the manager to incur a corresponding cost $E(e)$. In this model, to simplify, I assume e can take one of two values, that is $e \in \{\underline{e}, \bar{e}\}$ where $\bar{e} > \underline{e} > 0$. And the corresponding effort cost that the manager incurs can respectively be either $\underline{E}$ or $\bar{E}$ where $\bar{E} > \underline{E} > 0$. This setting ensures that the corresponding effort cost is increasing in the manager's effort, that is:

$$\frac{\partial E}{\partial e} > 0 \tag{13}$$

⁴ In Aghion et al. (2009) and Galasso and Simcoe (2011), the manager needs to determine whether to innovate or not. In the model presented here, the manager's concern is to exert effort.

Denote the firm-level TFP as a general function of the output Y , capital input K , labor input L , and intermediate input M such that the output and each input is a function of the manager's effort. Therefore, TFP can be written in a general form:

$$TFP = f(Y(e), K(e), L(e), M(e))$$

Without loss of generality and for the purpose of simplicity, assume TFP can be either 1 or 0.

If the CEO believes $\mathbb{C} = \bar{\mathbb{C}}$, the manager will make high effort $\bar{e}$ and incur a corresponding cost, $\bar{E}$, and TFP in period 1 is:

$$TFP_1 = \begin{cases} 1 & \text{with probability } \bar{p} \\ 0 & \text{with probability } 1 - \bar{p} \end{cases}$$

If the CEO believes $C = 0$, he will choose effort $\underline{e}$ and incur a corresponding cost, $\underline{E}$, and TFP in period 1 is equal to:

$$TFP_1 = \begin{cases} 1 & \text{with probability } \underline{p} \\ 0 & \text{with probability } 1 - \underline{p} \end{cases}$$

In their models, Aghion et al. (2009) and Galasso and Simcoe (2011) assume that if the manager does innovate, then the output will be realized in period 1. In this model, I assume that if the manager makes effort, the productivity is realized at the end of period 1 (i.e., at $t = 1$). This setting allows me to connect the managerial overconfidence and the firm-level TFP .

Assumption 1: $\underline{p} < \bar{p}$ to indicate that CEOs who choose low effort are less successful than those who choose high effort.

Assumption 1 implies that TFP is increasing in the manager's effort:

$$\frac{\partial TFP}{\partial e} > 0 \tag{14}$$

Assumption 1 is consistent with Holmstrom's (1979) result. Holmstrom (1979) shows that when the principal can only observe the firm's output and decides to pay the agent a proportion of the output, then the agent has to put forth higher effort to increase expected utility. This means that the agent will choose higher effort to achieve higher firm output because doing that also benefits himself, provided that the share of the output is sufficiently high so as to outweigh the cost of putting forth a high effort.

The model proceeds as follows: (1) First, the CEO determines whether to make effort e and pays the effort cost E . (2) Period 1 productivity TFP is realized. The market observes productivity and then updates its assessment of the project's ending cash flow. (3) The CEO compares his expected period 2 utility and his outside option to decide whether to leave the firm.

To continue, I make Assumptions 2 to 5 that are related to the CEO's compensation.

Assumption 2: The firm pays the manager based on the project's ending cash flow.

In their models, Aghion et al. (2009) and Galasso and Simcoe (2011) assume that the firm will pay the manager based on his ability. In this model, I assume that the firm will pay the manager based on the outcome (i.e. the ending cash flow) of the project. Assumption 2 is similar to the setting in the agency models in which the principal pays the agent based on the ultimate outcome or payoff of an action (Holmstrom, 1979).

Then I follow Aghion et al. (2009) and Galasso and Simcoe (2011) and make the following similar assumptions:

Assumption 3: if the manager chooses to leave the firm, what happens to his current job will not affect the compensation that he can get at the new job.

Assumption 4: the market only updates its assessment of the project's end of cash flow if the manager pays effort.⁵

As in Aghion et al. (2009) and Galasso and Simcoe (2011), I use backward induction to solve the model.

Assumptions 2 and 3 implies that if the manager goes to a new job, the new employer does not know what he did in his old firm and only pays him based on the market's prior beliefs about the project's end of cash flow. This means if the manager goes to a new firm, his expected utility is equal to:

$$\underline{w} = \frac{1}{2}(\bar{C} + 0) - s = \frac{1}{2}\bar{C} - s \quad (15)$$

⁵ In their models, Aghion et al. (2009) and Galasso and Simcoe (2011) assume that the market only updates its assessment on the manager's ability if the manager does innovate in period 1.

In Equation 15, $\frac{1}{2}\bar{C}$ is the expected salary that the manager can get if he relocates to a new firm and s is the switching cost. $\underline{w}$ represents the manager's outside option.

According to Assumption 4, if the CEO pays an effort and remains in the firm, the market will update its beliefs, following Bayes' rule. The CEO's expected wage in period 2 equals:

$$w_2(TFP_1) = \Pr(C = \bar{C}|TFP_1) \bar{C} + \Pr(C = 0|TFP_1) 0 = \Pr(C = \bar{C}|TFP_1) \bar{C}$$

If $TFP_1 = 1$, then:

$$w_2(TFP_1 = 1) = \Pr(C = \bar{C}|TFP_1 = 1) \bar{C} = \frac{\bar{p}}{\bar{p} + \underline{p}} \bar{C}$$

If $TFP_1 = 0$, then:

$$w_2(TFP_1 = 0) = \Pr(C = \bar{C}|TFP_1 = 0) \bar{C} = \frac{1 - \bar{p}}{1 - \bar{p} + 1 - \underline{p}} \bar{C} = \frac{1 - \bar{p}}{2 - \bar{p} - \underline{p}} \bar{C}$$

In alignment with Aghion et al. (2009) and Galasso and Simcoe (2011), I add the following assumption.

Assumption 5:⁶

$$\frac{\bar{p}}{\bar{p} + \underline{p}} \bar{C} > \frac{1}{2} \bar{C} - s = \underline{w} > \frac{1 - \bar{p}}{2 - \bar{p} - \underline{p}} \bar{C}$$

This assumption suggests that if the manager's first period productivity is low, he will leave the firm.

In period 1, the manager will choose to exert effort e if his expected utility function, $U(e > 0) - E$, exceeds or at least equals the utility from not exerting effort, $U(e = 0)$.

Assumption 4 suggests that if the manager does not make effort ($e = 0$) in period 1, the market will not have information to update its beliefs and only pays the manager based on its prior beliefs. Hence, the manager's expected utility in period 2 just equals to:

⁶ Because $\frac{\bar{p}}{\bar{p} + \underline{p}} > \frac{1}{2} > \frac{1 - \bar{p}}{2 - \bar{p} - \underline{p}}$ is true, one can always find at least a set of parameters $(\bar{p}, \underline{p}, \underline{w})$ that satisfies Assumption

$$U(e = 0) = \frac{1}{2}(\bar{C} + 0) = \frac{1}{2}\bar{C}$$

If the manager makes effort in period 1, the market will update its beliefs as the CEO's beliefs. Therefore, the manager's expected utility in period 2 will be:

$$\begin{aligned} U(e > 0) = & \left(\frac{1}{2}(1 + \Delta)\bar{p} + \frac{1}{2}(1 - \Delta)\underline{p} \right) \frac{\bar{p}}{\bar{p} + \underline{p}} \bar{C} \\ & + \left(\frac{1}{2}(1 + \Delta)(1 - \bar{p}) + \frac{1}{2}(1 - \Delta)(1 - \underline{p}) \right) \underline{w} \end{aligned} \quad (16)$$

Assumption 5 implies that if the manager remains in the same firm, the upper bound salary he can receive in period 2 is $\frac{\bar{p}}{\bar{p} + \underline{p}} \bar{C}$ and the lower bound salary he must receive in period 2 is $\underline{w} = \frac{1}{2}\bar{C} - s$.⁷ Therefore, the intuition of Equation 16 is that the manager's expected utility in period 2 is just equal to the weighted sum of the two possible salary levels that he can receive in period 2 and the weight of each salary level is just the probability that salary outcome can occur based on the manager's beliefs. In the right-hand side of Equation 16, the first term refers to the upper bound salary the manager can receive in period 2, $\frac{\bar{p}}{\bar{p} + \underline{p}} \bar{C}$, multiplied by the probability that salary outcome can happen. Similarly, the second term is the product of the lower bound salary the manager can receive in period 2, $\underline{w}$, and the probability that salary outcome can happen.

As noted previously, the manager only makes effort if:

$$U(e > 0) - E \geq U(e = 0) \quad (17)$$

Condition 17 only happens if the effort cost E is not too large. This means:⁸

$$\begin{aligned} E \leq E_{max} = & U(e > 0) - U(e = 0) \\ = & \frac{1}{2}\bar{p}\bar{C} + \left(1 - \frac{1}{2}\bar{p} - \frac{1}{2}\underline{p} \right) \underline{w} + \frac{1}{2}\Delta(\bar{p} - \underline{p}) \left(\frac{\bar{p}}{\bar{p} + \underline{p}} \bar{C} - \underline{w} \right) - \frac{1}{2}\bar{C} \end{aligned} \quad (18)$$

⁷ If he receives lower than $\underline{w}$, he will leave the firm.

⁸ Condition 18 is obtained through a process of algebra manipulation. Please see Appendix A for the process.

In Condition 18, E_{max} represents the maximum value of the manager's effort cost so that Condition 17 holds. Condition 18 implies that the manager's effort cost E is increasing in the overconfident factor Δ . This is because:⁹

$$\frac{\partial E}{\partial \Delta} = \frac{1}{2}(\bar{p} - \underline{p}) \left(\frac{\bar{p}}{\bar{p} + \underline{p}} \bar{c} - \underline{w} \right) > 0 \quad (19)$$

Use the chain rule for the left-hand side of Equation 19, one obtains:

$$\frac{\partial E}{\partial \Delta} = \frac{\partial E}{\partial e} \frac{\partial e}{\partial \Delta} > 0$$

And by Condition 13, $\frac{\partial E}{\partial e} > 0$, this means:

$$\frac{\partial e}{\partial \Delta} > 0 \quad (20)$$

Then take the derivative of the firm-level TFP with respect to the overconfident factor Δ and use the chain rule:

$$\frac{\partial TFP}{\partial \Delta} = \frac{\partial TFP}{\partial e} \frac{\partial e}{\partial \Delta}$$

By Condition 14, $\frac{\partial TFP}{\partial e} > 0$, and by Condition 20, $\frac{\partial e}{\partial \Delta} > 0$, therefore:

$$\frac{\partial TFP}{\partial \Delta} = \frac{\partial TFP}{\partial e} \frac{\partial e}{\partial \Delta} > 0 \quad (21)$$

The Result 21 predicts that:

Empirical Prediction: Overconfident managers are likely to achieve higher firm-level TFP.

The intuition behind the result is that overconfident managers often have stronger beliefs about the project's prospects. The beliefs induce them to exert higher effort, hence increase the probability that the project can be successful and consequently, the productivity increases. The result is

⁹ The right-hand side of Equation 19 is greater than 0 due to Assumptions 1 and 5.

consistent with theoretical findings and empirical evidence in existing literature. In their model, Gervais et al. (2011) point out that overconfident managers are more likely to make greater effort to learn about and improve the success rate of projects. Phua et al. (2018) show that overconfident CEOs have stronger beliefs about their firms' prospects. These managers make effort to spread the beliefs to other stakeholders such as the suppliers and the employees. Consequently, these stakeholders make longer commitments with the firms and eventually, might help improve firm productivity. In addition, Galasso and Simcoe (2011) and Hirshleifer et al. (2012) show that overconfident leaders are likely to make more innovative decisions which set the foundation up to boost the firm productivity.

In the next sections, I validate the prediction of the model by providing empirical evidence to support the result of the model.

4 Data and Methodology

4.1 Data

I collect data for a sample of S&P 1500 companies from 1992 to 2023. To measure overconfidence, I download executive compensation data from ExecuComp, which is available via Wharton Research Data Service (WRDS). In the ExecuComp database, the variable "SPIndexspcode" indicates whether a firm belongs to an S&P index. In the variable "SPIndexspcode", I choose the companies that are marked as "SP" which means the firm is in the S&P 500 index, or "MD" which means the firm is in the S&P Midcap index, or "SM" which means the firm is in the S&P Smallcap index. In addition, I focus on CEO overconfidence, hence, in the variable "AnnualCEOFlageoann", I select the observations that are identified as "CEO" which indicates the person is the CEO of the company in the year. I also obtained managerial characteristics such as gender and age from the ExecuComp data.

To measure managerial overconfidence, because I do not have the detailed dataset that Malmendier and Tate (2005) use in their study, I follow Campell et al. (2011) and Hirshleifer et al. (2012) to estimate the CEO's option moneyness and use it to classify an overconfident manager. The process of computing option moneyness requires the stock price at the end of the fiscal year which I have to download from Compustat. Furthermore, to calculate TFP and firm control variables, I also collect data from Compustat. Then I use the firm identifier (GVKEY) to match executive data from

ExecuComp with accounting data from Compustat. The calculation of TFP also requires price deflators which are obtained from Bureau of Economic Analysis - Table 1.1.9. Implicit Price Deflators for Gross Domestic Product.¹⁰ I follow previous studies and exclude financial firms by removing observations that have SIC codes ranging from 6000 to 6999 due to the difference in their accounting treatment, regulations and capital structure (Hirshleifer et al., 2012). In addition, to be consistent with prior studies, I also remove utilities firms that have SIC codes from 4900 to 4999 (Hirshleifer et al., 2012).

4.2 Variable Definitions

4.2.1 Total Factor Productivity

The dependent variable in this study is the firm-level total factor productivity (*TFP*). In economics, *TFP* is an overall measure of the effectiveness of the production process (İmrohoroglu and Tüzel 2014). To empirically measure *TFP*, in this paper, I use the production function approach and employ the Levinsohn and Petrin (LP) estimation. To justify the use of the LP approach, I first discuss the Ordinary Least Square (OLS) estimation to set up the foundation to see how the LP method can overcome the drawback of the OLS approach.

The Ordinary Least Square (OLS) Estimation

To calculate *TFP*, I start with an assumption that the production is in the simplest form of a Cobb-Douglas production function:

$$Y_{it} = A_{it} K_{it}^{\beta_K} L_{it}^{\beta_L} \quad (22)$$

where Y_{it} is the physical output of firm i in year t , K_{it} is the capital input of firm i in year t , L_{it} is the labor input of firm i in year t , and A_{it} is the Hicksian neutral efficiency level of firm i in year t (Hulten 2001; Van Beveren 2012).

Taking natural logs for both sides of Equation 22 will give:

$$\ln(Y_{it}) = \ln(A_{it}) + \ln(K_{it}^{\beta_K}) + \ln(L_{it}^{\beta_L})$$

¹⁰ U.S. Bureau of Economic Analysis, "[Table 1.1.9. Implicit Price Deflators for Gross Domestic Product](#)" (accessed Tuesday, October 22, 2024).

$$\ln(Y_{it}) = \ln(A_{it}) + \beta_K \ln(K_{it}) + \beta_L \ln(L_{it}) \quad (23)$$

Equation 23 can be rewritten in terms of a linear production function as follows:

$$\ln(Y_{it}) = \beta_0 + \beta_K \ln(K_{it}) + \beta_L \ln(L_{it}) + \varepsilon_{it} \quad (24)$$

where (Van Beveren 2012):

$$\ln(A_{it}) = \beta_0 + \varepsilon_{it}$$

In Equation 24, the OLS estimation will yield unbiased estimates if the zero conditional mean assumption holds, that is:

$$E(\varepsilon|K, L) = 0 \quad (25)$$

The zero conditional mean assumption says that there is no correlation between ε and the inputs K and L . If the zero conditional mean assumption holds, economists can use OLS to obtain consistent and unbiased estimates of the parameters β_K and β_L , and based on that, measure the firm-level *TFP*. Specifically, one can estimate *TFP* by first, regressing $\ln(Y_{it})$ on $\ln(K_{it})$ and $\ln(L_{it})$. Then *TFP* can be calculated as follows:

$$\widehat{TFP}_{it}^{OLS} = \ln(Y_{it}) - \hat{\beta}_K \ln(K_{it}) - \hat{\beta}_L \ln(L_{it}) \quad (26)$$

where $\hat{\beta}_K$ and $\hat{\beta}_L$ are estimated from the OLS estimation.

The Levinsohn and Petrin (LP) Estimation

One problem with the OLS estimation is that the zero conditional mean assumption 25 often does not hold in practice and this problem refers to the endogeneity issue. The endogeneity issue can arise, for example, firms with higher productivity require more inputs (Olley and Pakes 1996; Levinsohn and Petrin 2003), this implies $E(\varepsilon|K, L) \neq 0$. In this case, the OLS estimation will yield biased estimates of β_K and β_L and therefore, *TFP*.

To overcome the issue, Olley and Pakes (1996) split ε_{it} into two components: the firm productivity and the error term, then Equation 24 can be rewritten as follows:

$$\ln(Y_{it}) = \beta_0 + \beta_K \ln(K_{it}) + \beta_L \ln(L_{it}) + \theta_{it} + \mu_{it} \quad (27)$$

In Equation 27, θ_{it} is the productivity and μ_{it} is the error term (Olley and Pakes, 1996). In this case, there is a correlation between the inputs and the productivity θ_{it} . To control for this issue, they employ the investment proxy and write the investment as a function of K_{it} and θ_{it} :

$$I_{it} = I_{it}(K_{it}, \theta_{it})$$

Then they use a semiparametric procedure to estimate the parameters (i.e., β_K and β_L) of the production function, and they show that the estimates obtained from their estimation algorithm are consistent and unbiased. However, as noted by Levinsohn and Petrin (2003), the use of the investment proxy might lead to some concerns. First, the proxy does not work if firms do not report positive investment. Second, in case if the adjustment costs are too large to respond to a productivity shock, the use of the investment proxy does not eliminate well the correlation between the inputs and the unobserved factors. Therefore, Levinsohn and Petrin (2003) propose to use intermediate inputs such as materials or electricity as the alternatives of the investment proxy. They also emphasize that intermediate inputs are often not state variables and therefore, offer a simple connection between the estimation procedure and the economy theory. In addition, under some specific conditions, the intermediate inputs can be good proxies for the productivity shock. In this research, I follow the method in Levinsohn and Petrin (2003) and write the production function as follows:

$$\ln(Y_{it}) = \beta_0 + \beta_K \ln(K_{it}) + \beta_L \ln(L_{it}) + \beta_M \ln(M_{it}) + \theta_{it} + \mu_{it} \quad (28)$$

In Equation 28, the subscripts i and t denote firm i and year t , respectively. Following prior studies (Craig and Harris 1973; Levinsohn and Petrin 2003; Petrin et al. 2004; De Loecker 2011; Gal 2013; İmrohoroglu and Tüzel 2014), I use net sales deflated by the price deflator to be a proxy for the physical output Y , property plant and equipment deflated by the price deflator to measure the capital input K , number of employees to represent the labor input L . Finally, M is the materials and proxied by the firm's inventory deflated by a GDP deflator. Then I estimate the firm-level TFP in a log form by:

$$\widehat{TFP}_{it}^{LP} = \ln(Y_{it}) - \hat{\beta}_K \ln(K_{it}) - \hat{\beta}_L \ln(L_{it}) - \hat{\beta}_M \ln(M_{it}) \quad (29)$$

where $\hat{\beta}_K$, $\hat{\beta}_L$, and $\hat{\beta}_M$ are estimated from the semiparametric procedure in Levinsohn and Petrin (2003).¹¹

4.2.2 CEO Overconfidence

¹¹ Please see Appendix B for the procedure of the LP estimation.

To evaluate whether a manager is overconfident or not, Malmendier and Tate (2005) propose three measures, namely Holder67, Longholder, and Net Buyer. Due to data limitations, the Longholder metric cannot be constructed using the ExecuComp database. Therefore, in this paper, I focus on the other two metrics, Holder67 and Net Buyer. The intuition behind Holder67 is that a risk-averse and undiversified manager tends to exercise her firm's stock options early if the option moneyness exceeds a rational benchmark (Hall and Murphy, 2002). If a manager fails to exercise the option right after the vesting period when the option is at least 67% in the money, it is an indicator that the manager is overconfident (Malmendier and Tate, 2005). I follow the idea of Holder67 and employ the strategy that Campell et al. (2011) and Hirshleifer et al. (2012) used to estimate the top manager's option moneyness. Specifically, first, I calculated the value per option which equals the total value of the unexercised exercisable options divided by the number of unexercised exercisable options. Second, I estimated the exercise price per option by subtracting the value per option from the stock price at the end of the fiscal year. Last, the option moneyness is calculated by taking the value per option divided by the exercise price per option. I then follow Hirshleifer et al. (2012) and track the entire tenure of a manager and if during his tenure, there is at least one time that the option moneyness is greater than or equal to 67 percent, the variable *OVERCONFID67* will get the value of 1 starting from the first time the moneyness is at least 67 percent and 0 otherwise.¹² I use the *OVERCONFID67* variable for my main results.

For the robustness check, I also use the Net Buyer measure proposed by Malmendier and Tate (2005). To be precise, the *NET BUYER* variable in this study will get the value of 1 if during the first five years of a manager's tenure, he was more often a net buyer than a net seller. To construct this variable, I require a CEO needs to appear at least 6 years otherwise I will treat the CEO as missing observation. The first year is the base year, and the next five years are to determine the net buying behavior of the manager.

4.2.3 Control Variables

In alignment with prior studies (Malmendier and Tate 2005, Hirshleifer et al. 2012, İmrohoroglu and Tüzel 2014), I include control variables that might be related to both CEO overconfidence and

¹² In untabulated results when I constructed the *OVERCONFID67* based on the requirement that the moneyness option has to be greater than or equal to 67 percent at least twice, the main findings still hold.

firm-level TFP. First, I add CEO characteristics. *CEO AGE* is the age of the CEO in concern. *CEO OWNERSHIP* is the ownership proportion that the CEO holds in his company and ownership data is obtained from ExecuComp. *CEO FEMALE* is a binary variable, equals 1 if the CEO is female and 0 otherwise. Second, I also include firm specifications. *ROA* is the return on assets and equals net income divided by total assets. *TOBIN'S Q* is measured by taking the total market value of liabilities and equity divided by the book value of total assets. Because the market value of liabilities is unavailable, the book value is used as a proxy instead. *SIZE* is the natural logarithm of total assets. *LEVERAGE* is defined as total liabilities scaled by total assets. *CASH* is calculated as cash and cash equivalents divided by total assets. Finally, *R&D* equals research and development expense divided by total assets. All continuous variables are winsorized at 1 and 99 percent. Table 1 summarizes the definitions and calculations of all the variables.

[INSERT TABLE 1 HERE]

4.2.4 Descriptive Statistics and Correlation Analysis

Table 2 presents the descriptive statistics (Panel A) and compares the means of all variables by *OVERCONFID67* (Panel B) and *NET BUYER* (Panel C). In Panel A, the average value of TFP^{LP} is 1.232. The values are consistent with the ones estimated in İmrohoroglu and Tüzel (2014) in which the authors also used Compustat data to calculate the firm-level *TFP*. In the sample, 61.7 percent of managers are overconfident if using the option moneyness of at least 67 percent. Alternatively, when using the net buying behavior in the first five years of a manager's tenure, 75 percent of CEOs are overconfident. These statistics are consistent with previous studies (Goel and Thakor 2008, Campell et al. 2011, Hirshleifer et al. 2012) where overconfidence is a dominant characteristic among CEOs. A typical manager is 57 years old and holds 2.7 percent of his firm's stock. In the sample, only 3.5 percent of the top leaders are female. On average, a firm generates an annual return on assets of 5.4 percent, has a market value that is more than double its book value, uses debts to finance 54.9 percent of its assets, holds 10.7 percent of total assets as cash, and spends 4.6 percent of its assets on research and development expenses.

[INSERT TABLE 2 HERE]

In Panel B, when the measure of CEO overconfidence is *OVERCONFID67*, overconfident managers seem to create more productivity as evidenced by higher TFP^{LP} , be older, hold less the ownership of the firms, and less likely to be a female than their non-overconfident counterparts.

In addition, overconfident CEOs tend to be associated with higher returns on assets, greater firm value, more cash holdings, more spending on research and development, smaller firm size, and use less debts in financing decisions. These results are consistent with Hirshleifer et al. (2012).

In Panel C, when the measure of managerial overconfidence is *NET BUYER*, overconfident CEOs also have higher *TFP*. A net buying manager tends to be younger, hold less the firm's shares and more likely to be a female than their non-overconfident peers. Beyond that, a net buyer is also associated with higher return on assets, larger firm size, more leverage, lower firm value, less cash holdings, and less spending on research and development activities. These results seem to be consistent with Malmendier and Tate (2005) although they use a different sample of U.S. firms.

Table 3 presents the correlation coefficients between variables. The two metrics of managerial overconfidence seem to be positively correlated with each other although the magnitude of the relationship might be small. The correlation coefficient between *OVERCONFID67* and *NET BUYER* is 0.024 and significant. Malmendier and Tate (2005) use a different sample and the correlation coefficient between the two measures in their paper is 0.06. Given a small coefficient between *OVERCONFID67* and *NET BUYER*, one can expect that although the two metrics try to measure managerial overconfidence, they still can capture some different aspects of managers' expectations related to firms' prospects. One potential difference is *OVERCONFID67* might reflect a short-term expectation of a manager on the firm's market value while *NET BUYER* might represent a long-term perspective of a manager on the firm's prospects. This might be true because *NET BUYER* is an indicator of a net buying behavior of firm shares that occurs over a long period (i.e., 5 years).

[INSERT TABLE 3 HERE]

The results in Table 3 also show that managerial overconfidence is positively correlated with the firm productivity. When the measure of overconfidence is *OVERCONFID67*, it is positively correlated with TFP^{LP} . The correlation coefficient between *OVERCONFID67* and TFP^{LP} is 0.053 and statistically significant. When managerial overconfidence is proxied by *NET BUYER*, the correlation coefficient between the variable and *TFP* is again positively significant. The coefficient between net buyers and firms' LP-estimated productivity is 0.083.

4.3 Model Specification

To test the empirical prediction of the model in Section 3, I employ the following regression:¹³

$$\begin{aligned}
TFP_{it}^{LP} = & \beta_0 + \beta_1 OVERCONFIDENCE_{it} + \beta_2 CEO\ AGE_{it} \\
& + \beta_3 CEO\ OWNERSHIP_{it} + \beta_4 CEO\ FEMALE_{it} + \beta_5 ROA_{it} \\
& + \beta_6 TOBIN'S\ Q_{it} + \beta_7 SIZE_{it} + \beta_8 LEVERAGE_{it} + \beta_9 CASH_{it} \\
& + \beta_{10} R\&D_{it} + \gamma_i + \delta_j + \theta_t + \varepsilon_{it}
\end{aligned} \tag{30}$$

In Equation 30, the subscripts i and t denote firm i and year t , respectively.¹⁴ TFP_{it}^{LP} measures the firm-level TFP which is estimated using the LP approach. *OVERCONFIDENCE* is a measure of managerial overconfidence and can be proxied by *OVERCONFID67* or *NET BUYER*. *CEO AGE*, *CEO OWNERSHIP*, and *CEO FEMALE* represent CEO characteristics. Firm specifications include *ROA*, *TOBIN'S Q*, *SIZE*, *LEVERAGE*, *CASH*, and *R&D*. In Equation 30, I also include the firm fixed effect γ_i , the industry fixed effect δ_j , and the year fixed effect θ_t . These fixed effects are to deal with the potential issue of unobserved omitted variables. In addition, when I run Regression 30, I follow prior studies (Petersen 2008; Thompson 2011) and calculate heteroskedasticity robust standard errors clustered by firm and year to control for the correlation of the residuals across firms and across time.

5 Empirical Evidence

5.1 Option-Based Measure of Overconfidence and TFP LP

Table 4 reports the results when I regress TFP_{it}^{LP} on the option-based measure of overconfidence, *OVERCONFID67*. Column 1 is the result when I do not include *R&D*, and the firm fixed effect. Column 2 is the result when I add the firm fixed effect but still exclude *R&D*. Column 3 presents the result when I control for *R&D*. Because there are many missing values in *R&D*, the number of observations drop significantly in Column 3, compared to Columns 1 and 2. To deal with these missing values in *R&D*, I use the most two common approaches that are often used in the literature. The first one is to replace missing *R&D* observations by zero and the second one is to replace them

¹³ In all the regressions in this paper, the Variance Inflation Factors (VIFs) are small, ranging from 1.01 to 1.47, suggesting that the multicollinearity is not a serious concern.

¹⁴ In untabulated results, to avoid reverse causality, I also regress the dependent variable on the one-year lags of all the independent variables and the results are still consistent with the main findings.

by the industry averages (Koh and Reeb, 2015). After replacing missing *R&D* values, I rerun the regression. Columns 4 and 5 respectively report the regression results of the two approaches.

[INSERT TABLE 4 HERE]

From Columns 1 to 5, the estimated coefficients of *OVERCONFID67* are all positive and significant, indicating a positive effect of managerial overconfidence on the firm productivity. In Table 4, Column 2 (i.e., I do not add *R&D*) and Column 4 (i.e., I replace missing *R&D* by zero) give the same coefficient estimates of *OVERCONFID67*, 0.025. In Column 3 when I include *R&D*, the estimated coefficient slightly decreases to 0.023. And when I replace missing *R&D* by the industry averages, the coefficient estimate slightly increases to 0.028.

In Column 4, the estimated coefficient of *OVERCONFID67* is 0.025 and significant, implying that an overconfident leader can increase the firm productivity by $100 \times (e^{0.025} - 1) = 2.53\%$, compared to a non-overconfident manager. In Column 5, the increase in the firm productivity is 2.84% due to the impact of managerial overconfidence.

Regarding control variables, both CEO and firm characteristics can affect firm productivity. An increase in CEO age is associated with a decrease whereas an increase in CEO ownership is related to a boost in firm productivity (Columns 1 to 5). A female CEO seems to be associated with a lower firm productivity. Firm's return on assets, firm value, firm size, and cash holding can increase the productivity of the firm. The results in Table 4 do not show a clear impact of firm leverage on firm productivity.

5.2 Net Buying Behavior-Based Measure of Overconfidence and TFP LP

To confirm if the results still hold for the net buying behavior-based measure of managerial overconfidence, I also regress TFP^{LP} on *NET BUYER* and Table 5 presents the corresponding results.

[INSERT TABLE 5 HERE]

All the coefficient estimates of *NET BUYER* are positive and statistically significant, confirming the findings in the previous sections. Column 2 (i.e., I do not control for *R&D*) and Column 4 (i.e., I replace missing *R&D* by zero) give stable and significant coefficient estimates of *NET BUYER* and both the coefficients are 0.056. The result in Column 4 indicates that compared to a non-net buyer, a net buying CEO increase the firm productivity by $100 \times (e^{0.056} - 1) = 5.76\%$.

In Column 3, when I include *R&D* in the regression, the coefficient estimate of *NET BUYER* is 0.028. In Column 5, when I replace missing *R&D* by the industry averages, the corresponding estimated coefficient is 0.061. Both the coefficients are statistically significant. Recall that in Table 4, I do not see a significant difference in the results between Columns 3 and 5. However, in Table 5, there is a discrepancy in the coefficient estimates of the managerial overconfidence in the two columns. One potential explanation is that the *NET BUYER* variable has more missing values, compared to the *OVERCONFID67* variable. In addition, the *R&D* variable has already had significant missing values. Therefore, when I include both the *NET BUYER* and the *R&D* variables in the regression in Column 3, the number of observations drop further and consequently, affecting the estimated coefficient of *NET BUYER*. In contrast, when I replace missing *R&D* by the industry averages as in Column 5, the impact of the net buyers significantly increases. In Column 5, the coefficient of 0.061 suggests that an overconfident manager can increase the firm productivity by $100 \times (e^{0.061} - 1) = 6.29\%$, compared to his non-overconfident peers.

Regarding control variables, robust evidence is found for CEO age, CEO ownership, firm's return on assets, firm value, firm size, and cash holding which are similar to the ones observed in Table 4. An increase in CEO ownership is expected to boost the firm productivity up. A higher firm return on asset, a greater firm value, a larger firm size and amount of cash holding are also associated with an increase in the firm productivity.

5.3 CEO Change Analysis

So far, all the empirical evidence confirm that overconfident managers can bring positive impact on firm productivity. Although I tried different model specifications and control for CEO characteristics, firm characteristics and firm fixed effect, there might still be potential omitted variables that can affect both the managerial overconfidence and firm productivity. Therefore, the risk of endogeneity still exists. To deal with this issue, I follow prior studies (Galasso and Simcoe 2011; Kim et al. 2016; Chen et al. 2020) and conduct CEO change analysis. Specifically, I create subsamples for firms that within the firms, there are switches from non-overconfident to overconfident CEOs or vice versa. To do that, I create two separate subsamples that corresponds to the two measures of overconfidence, one for *OVERCONFID67* and the other for *NET BUYER*. Then I rerun the Regression 30 for the two subsamples. Table 6 presents the results for the subsample of *OVERCONFID67*.

[INSERT TABLE 6 HERE]

In Table 6, from Columns 1 to 5, all the coefficient estimates of *OVERCONFID67* are positive and significant. In Column 1, the estimated coefficient is 0.033, implying that compared to a non-overconfident manager, an overconfident one can increase firm productivity by $100 \times (e^{0.033} - 1) = 3.36\%$. In Column 4, the increase in firm productivity as the result of appointing an overconfident leader is 4.60%.

[INSERT TABLE 7 HERE]

Similarly, Table 7 shows the results for the subsample of *NET BUYER*. Again, all the estimated coefficients of *NET BUYER* from Columns 1 to 5 confirm the main findings. The coefficients are all positive and significant. The coefficient estimate of *NET BUYER* in Column 5 is 0.058, indicating that a net-buying leader is expected to boost the firm productivity up by $100 \times (e^{0.058} - 1) = 5.97\%$.

6 Conclusions

The existing literature has provided empirical evidence to show that CEO overconfidence can affect the firms' production process. In this paper, I add to the current literature by investigating the relationship between managerial overconfidence and the firm-level TFP. I modify the models in Aghion et al. (2009) and Galasso and Simcoe (2011) to explain how CEO overconfidence can have impacts on firm productivity. The result of the model show that overconfident leaders have stronger beliefs about the firms' prospects. Therefore, the beliefs inspire them to exert efforts to realize their vision. And by putting higher efforts on their firms, they increase the firm productivity. I also validate the prediction of the model by providing empirical evidence. The evidence shows that overconfident CEOs are likely to achieve higher firm productivity. The results are robust across different model specifications, fixed effects, different measures of firm-level TFP and managerial overconfidence, and also within-firm analysis.

References

Abramovitz, M. (1956). Resource and output trends in the United States since 1870. In *Resource and Output Trends in the United States since 1870* (pp. 1-23). NBER.

- Akerberg, D., Caves, K., & Frazer, G. (2006). Structural identification of production functions. *Munich Personal RePEc Archive (MPRA) Paper 38349*.
- Aghion, P., Van Reenen, J., & Zingales, L. (2009). Innovation and institutional ownership. *NBER Working Paper 14769*.
- Barton, G. T., & Cooper, M. R. (1948). Relation of agricultural production to inputs. *The Review of Economics and Statistics*, 30(2), 117-126.
- Brynjolfsson, E., & Hitt, L. M. (2003). Computing productivity: Firm-level evidence. *Review of Economics and Statistics*, 85(4), 793-808.
- Campbell, T. C., Gallmeyer, M., Johnson, S. A., Rutherford, J., & Stanley, B. W. (2011). CEO optimism and forced turnover. *Journal of Financial Economics*, 101(3), 695-712.
- Chen, Y. R., Ho, K. Y., & Yeh, C. W. (2020). CEO overconfidence and corporate cash holdings. *Journal of Corporate Finance*, 62, 101577.
- Cobb, C. W., & Douglas, P. H. (1928). A theory of production. *The American Economic Review*, 18(1), 139-165.
- Copeland, M. A. (1937). Concepts of national income. In *Studies in Income and Wealth, Volume 1* (pp. 2-63). NBER.
- Copeland, M. A., & Martin, E. M. (1938). The correction of wealth and income estimates for price changes. In *Studies in Income and Wealth, Volume 2* (pp. 83-134). NBER.
- Craig, C. E., & Harris, R. C. (1973). Total productivity measurement at the firm level. *Sloan Management Review (pre-1986)*, 14(3), 13.
- De Loecker, J. (2011). Product differentiation, multiproduct firms, and estimating the impact of trade liberalization on productivity. *Econometrica*, 79(5), 1407-1451.
- Fabricant, S. (1954). Economic progress and economic change. In *Economic Progress and Economic Change*. NBER.
- Ferris, S. P., Jayaraman, N., & Sabherwal, S. (2013). CEO overconfidence and international merger and acquisition activity. *Journal of Financial and Quantitative Analysis*, 48(1), 137-164.
- Gal, P. N. (2013). Measuring total factor productivity at the firm level using OECD-ORBIS.

- Galasso, A., & Simcoe, T. S. (2011). CEO overconfidence and innovation. *Management Science*, 57(8), 1469-1484.
- Gervais, S., Heaton, J. B., & Odean, T. (2011). Overconfidence, compensation contracts, and capital budgeting. *The Journal of Finance*, 66(5), 1735-1777.
- Goel, A. M., & Thakor, A. V. (2008). Overconfidence, CEO selection, and corporate governance. *The Journal of Finance*, 63(6), 2737-2784.
- Griliches, Z. (1996). The discovery of the residual: a historical note. *Journal of Economic Literature*, 34(3), 1324-1330.
- Hall, B. J., & Murphy, K. J. (2002). Stock options for undiversified executives. *Journal of Accounting and Economics*, 33(1), 3-42.
- Heaton, J. B. (2002). Managerial optimism and corporate finance. *Financial Management*, 31(2), 33-45.
- Hirshleifer, D., Low, A., & Teoh, S. H. (2012). Are overconfident CEOs better innovators?. *The Journal of Finance*, 67(4), 1457-1498.
- Holmstrom, B. (1979). Moral hazard and observability. *The Bell Journal of Economics*, 10(1), 74-91.
- Holmstrom, B. (1982). Managerial incentive problems: A dynamic perspective. *Essays in Economics and Management in honor of Lars Wahlbeck*, Helsinki: Swedish School of Economics
- Hulten, C. R. (2001). Total factor productivity: a short biography. In *New Developments in Productivity Analysis* (pp. 1-54). University of Chicago Press.
- İmrohoroğlu, A., & Tüzel, Ş. (2014). Firm-level productivity, risk, and return. *Management Science*, 60(8), 2073-2090.
- Javorcik, B. S. (2004). Does foreign direct investment increase the productivity of domestic firms? In search of spillovers through backward linkages. *American Economic Review*, 94(3), 605-627.
- Kendrick, J. W. (1955). Productivity. In *Government in Economic Life* (pp. 44-47). NBER.
- Kendrick, J. W. (1956). Productivity trends: capital and labor. NBER.

- Kim, J. B., Wang, Z., & Zhang, L. (2016). CEO overconfidence and stock price crash risk. *Contemporary Accounting Research*, 33(4), 1720-1749.
- Koh, P. S., & Reeb, D. M. (2015). Missing R&D. *Journal of Accounting and Economics*, 60(1), 73-94.
- Levinsohn, J., & Petrin, A. (2003). Estimating production functions using inputs to control for unobservables. *The Review of Economic Studies*, 70(2), 317-341.
- Malmendier, U., & Tate, G. (2005). CEO overconfidence and corporate investment. *The Journal of Finance*, 60(6), 2661-2700.
- Malmendier, U., & Tate, G. (2008). Who makes acquisitions? CEO overconfidence and the market's reaction. *Journal of Financial Economics*, 89(1), 20-43.
- Nagle, F. (2019). Open source software and firm productivity. *Management Science*, 65(3), 1191-1215.
- Olley, S., & Pakes, A. (1996). The dynamics of productivity in the telecommunications equipment industry. *Econometrica*, 64(6), 1263-1297.
- Parrotta, P., Pozzoli, D., & Pytlikova, M. (2014). Labor diversity and firm productivity. *European Economic Review*, 66, 144-179.
- Petersen, M. A. (2008). Estimating standard errors in finance panel data sets: Comparing approaches. *The Review of Financial Studies*, 22(1), 435-480.
- Petrin, A., Poi, B. P., & Levinsohn, J. (2004). Production function estimation in Stata using inputs to control for unobservables. *The Stata Journal*, 4(2), 113-123.
- Phua, K., Tham, T. M., & Wei, C. (2018). Are overconfident CEOs better leaders? Evidence from stakeholder commitments. *Journal of Financial Economics*, 127(3), 519-545.
- Ruttan, V. W. (1954). Technological progress in the meatpacking industry, 1919-1947.
- Ruttan, V. W. (1956). The Contribution of Technological Progress to Farm Output: 1950-75. *The Review of Economics and Statistics*, 38(1), 61-69.
- Schmookler, J. (1952). The changing efficiency of the American economy, 1869-1938. *The Review of Economics and Statistics*, 214-231.

- Solow, R. M. (1957). Technical change and the aggregate production function. *The Review of Economics and Statistics*, 39(3), 312-320.
- Stigler, G. J. (1947). Trends in output. *Trends in Output and Employment*, 6-31.
- Thompson, S. B. (2011). Simple formulas for standard errors that cluster by both firm and time. *Journal of Financial Economics*, 99(1), 1-10.
- Tinbergen, J. (1942). Zur theorie der langfristigen wirtschaftsentwicklung. *Weltwirtschaftliches Archiv*, 511-549.
- Van Beveren, I. (2012). Total factor productivity estimation: A practical review. *Journal of Economic Surveys*, 26(1), 98-128.
- Wolff, E. N. (1996). The productivity slowdown: the culprit at last? Follow-up on Hulten and Wolff. *The American Economic Review*, 86(5), 1239-1252.

Table 1: Variable Definitions

This table presents the detailed definitions and calculations of all variables.

Abbreviated Variable Name	Full Written Name	Description
Dependent Variable		
<i>TFP^{LP}</i>	<i>Levinsohn and Petrin (LP) Estimation-based Total Factor Productivity</i>	Y is the firm output and proxied by net sales deflated by the price deflator. K is the capital input and proxied by property plan and equipment deflated by the price deflator. L is the labor input and proxied by number of employees. M are the materials and proxied by the firm's inventory deflated by a GDP deflator. The production function is: $\ln(Y_{it}) = \beta_0 + \beta_K \ln(K_{it}) + \beta_L \ln(L_{it}) + \beta_M \ln(M_{it}) + \theta_{it} + \mu_{it}$ Estimate $\hat{\beta}_K$, $\hat{\beta}_L$, and $\hat{\beta}_M$ by using the semiparametric procedure in Levinsohn and Petrin (2003). Then estimate TFP_{it}^{LP} as follows: $\widehat{TFP}_{it}^{LP} = \ln(Y_{it}) - \hat{\beta}_K \ln(K_{it}) - \hat{\beta}_L \ln(L_{it}) - \hat{\beta}_M \ln(M_{it})$
Independent Variables		
<i>OVERCONFID67</i>	<i>CEO Overconfidence</i>	If during a CEO's tenure, there is at least one time that the option moneyness is greater than or equal to 67 percent, this variable will get the value of 1 starting from the first time the moneyness is at least 67 percent and 0 otherwise.
<i>NET BUYER</i>	<i>CEO Net Buyer</i>	A binary variable gets the value of 1 if during the first five years of a manager's tenure, he was more often a net buyer than a net seller.
Control Variables		
<i>CEO AGE</i>	<i>The age of the CEO</i>	CEO age is obtained from ExecuComp.
<i>CEO OWNERSHIP</i>	<i>CEO Ownership</i>	The ownership proportion that the CEO is holding in his company. Data is obtained from ExecuComp.
<i>CEO FEMALE</i>	<i>CEO Female</i>	A binary variable equals 1 if the CEO is female and 0 otherwise.
<i>ROA</i>	<i>Return on Assets</i>	Equals net income divided by total assets
<i>TOBIN'S Q</i>	<i>Tobin's Q</i>	Equals the total market value of liabilities and equity divided by the book value of total assets. Because the market value of liabilities is unavailable, the book value is used as an alternative.
<i>SIZE</i>	<i>Firm Size</i>	Equals the natural logarithm of total assets

<i>LEVERAGE</i>	<i>Firm Leverage</i>	Equals total liabilities scaled by total assets
<i>CASH</i>	<i>Cash</i>	Equals cash and cash equivalents divided by total assets
<i>R&D</i>	<i>Research and Development Expense</i>	Equals research and development expense divided by total assets

Table 2: Descriptive Statistics

This table presents the descriptive statistics of all variables (Panel A) and the t -tests of all variables by *OVERCONFID67* (Panel B) and *NET BUYER* (Panel C). All variables are defined in Table 1.

Panel A: Descriptive Statistics						
Variables	Obs	Mean	Std. dev.	P1	P50	P99
<i>TFP^{LP}</i>	18,881	1.232	0.648	-2.226	1.175	5.576
<i>OVERCONFID67</i>	19,478	0.617	0.486	0.000	1.000	1.000
<i>NET BUYER</i>	17,329	0.750	0.433	0.000	1.000	1.000
<i>CEO AGE</i>	22,138	56.658	7.449	27.000	57.000	93.000
<i>CEO OWNERSHIP</i>	17,801	0.027	0.063	0.000	0.004	0.359
<i>CEO FEMALE</i>	22,536	0.035	0.183	0.000	0.000	1.000
<i>ROA</i>	22,460	0.054	0.094	-0.412	0.059	0.289
<i>TOBIN'S Q</i>	20,162	2.310	1.645	0.783	1.776	10.305
<i>SIZE</i>	22,465	8.002	1.612	4.344	7.909	12.192
<i>LEVERAGE</i>	22,412	0.549	0.224	0.095	0.547	1.272
<i>CASH</i>	22,127	0.107	0.108	0.001	0.072	0.535
<i>R&D</i>	14,719	0.046	0.064	0.000	0.021	0.355

Panel B: t-tests by <i>OVERCONFID67</i>					
Variables	<i>OVERCONFID67</i> = 1		<i>OVERCONFID67</i> = 0		Mean Difference
	Mean	No. of Observations	Mean	No. of Observations	
<i>TFP^{LP}</i>	1.260	10,095	1.189	6,276	0.071***
<i>CEO AGE</i>	57.369	11,855	56.728	7,278	0.641***
<i>CEO OWNERSHIP</i>	0.026	9,910	0.029	5,888	-0.003***
<i>CEO FEMALE</i>	0.027	12,012	0.043	7,466	-0.015***
<i>ROA</i>	0.064	12,011	0.044	7,431	0.020***
<i>TOBIN'S Q</i>	2.578	11,067	1.985	6,700	0.593***
<i>SIZE</i>	7.964	12,012	8.164	7,433	-0.200***
<i>LEVERAGE</i>	0.532	11,981	0.570	7,416	-0.038***
<i>CASH</i>	0.110	11,833	0.104	7,317	0.007***
<i>R&D</i>	0.048	7,942	0.038	4,708	0.009***

Panel C: t-tests by <i>NET BUYER</i>					
Variables	<i>NET BUYER</i> = 1		<i>NET BUYER</i> = 0		Mean Difference
	Mean	No. of Observations	Mean	No. of Observations	
<i>TFP^{LP}</i>	1.260	11,320	1.131	3,365	0.129***
<i>CEO AGE</i>	56.351	12,850	58.162	4,303	-1.811***
<i>CEO OWNERSHIP</i>	0.019	10,320	0.062	3,798	-0.043***
<i>CEO FEMALE</i>	0.038	12,989	0.009	4,340	0.029***
<i>ROA</i>	0.060	12,987	0.049	4,338	0.010***
<i>TOBIN'S Q</i>	2.273	12,046	2.463	3,911	-0.189***
<i>SIZE</i>	8.184	12,987	7.356	4,338	0.828***
<i>LEVERAGE</i>	0.555	12,965	0.489	4,326	0.066***
<i>CASH</i>	0.102	12,767	0.123	4,293	-0.021***
<i>R&D</i>	0.044	8,572	0.053	2,687	-0.009***

***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 3: Correlation Analysis

This table presents the correlation coefficients between variables. All variables are defined in Table 1.

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
<i>(1) TFP^{LP}</i>	1.000						
<i>(2) OVERCONFID67</i>	0.053***	1.000					
<i>(3) NET BUYER</i>	0.083***	0.024***	1.000				
<i>(4) CEO AGE</i>	-0.065***	0.042***	-0.105***	1.000			
<i>(5) CEO OWNERSHIP</i>	0.002	-0.023***	-0.286***	0.128***	1.000		
<i>(6) CEO FEMALE</i>	0.002	-0.042***	0.073***	-0.035***	-0.016**	1.000	
<i>(7) ROA</i>	0.166***	0.104***	0.050***	0.045***	0.038***	0.014**	1.000
<i>(8) TOBIN'S Q</i>	0.075***	0.171***	-0.050***	-0.073***	0.061***	0.019***	0.257***
<i>(9) SIZE</i>	0.051***	-0.061***	0.225***	0.115***	-0.235***	0.035***	0.077***
<i>(10) LEVERAGE</i>	0.009	-0.083***	0.132***	0.014**	-0.179***	0.018***	-0.175***
<i>(11) CASH</i>	0.104***	0.031***	-0.083***	-0.091***	0.096***	0.030***	-0.034***
<i>(12) R&D</i>	0.079***	0.074***	-0.060***	-0.143***	0.067***	-0.018**	-0.328***
Variables	(8)	(9)	(10)	(11)	(12)		
<i>(8) TOBIN'S Q</i>	1.000						
<i>(9) SIZE</i>	-0.182***	1.000					
<i>(10) LEVERAGE</i>	-0.088***	0.359***	1.000				
<i>(11) CASH</i>	0.341***	-0.295***	-0.244***	1.000			
<i>(12) R&D</i>	0.368***	-0.276***	-0.170***	0.486***	1.000		

***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 4: Option-Based Measure of Overconfidence and TFP LP

This table presents the results when regressing TFP^{LP} on $OVERCONFID67$ along with control variables. All variables are defined in Table 1.

VARIABLE	Dependent Variable: TFP^{LP}				
	(1)	(2)	(3)	(4)	(5)
<i>OVERCONFID67</i>	0.064*** (0.010)	0.025*** (0.007)	0.023*** (0.008)	0.025*** (0.007)	0.028*** (0.007)
<i>CEO AGE</i>	-0.006*** (0.001)	-0.003*** (0.000)	-0.003*** (0.000)	-0.003*** (0.000)	-0.003*** (0.000)
<i>CEO OWNERSHIP</i>	0.380*** (0.086)	0.217*** (0.072)	0.368*** (0.093)	0.216*** (0.072)	0.222*** (0.075)
<i>CEO FEMALE</i>	-0.064*** (0.024)	-0.025** (0.012)	-0.016 (0.013)	-0.025** (0.012)	-0.025** (0.012)
<i>ROA</i>	0.981*** (0.064)	1.225*** (0.051)	1.102*** (0.063)	1.221*** (0.051)	1.170*** (0.052)
<i>TOBIN'S Q</i>	0.023*** (0.004)	0.013*** (0.003)	0.016*** (0.003)	0.013*** (0.003)	0.014*** (0.003)
<i>SIZE</i>	0.056*** (0.003)	0.085*** (0.006)	0.080*** (0.008)	0.085*** (0.006)	0.087*** (0.007)
<i>LEVERAGE</i>	0.189*** (0.026)	0.037 (0.023)	0.014 (0.028)	0.037 (0.023)	0.034 (0.024)
<i>CASH</i>	0.759*** (0.054)	0.114*** (0.037)	0.098** (0.043)	0.115*** (0.037)	0.142*** (0.038)
<i>R&D</i>			-0.228 (0.199)		
<i>R&D (missing values replaced by 0)</i>				-0.064 (0.192)	
<i>R&D (missing values replaced by industry averages)</i>					-0.047 (0.182)
Constant	0.800*** (0.046)	0.600*** (0.057)	0.621*** (0.075)	0.608*** (0.060)	0.597*** (0.062)
Firm fixed effect	No	Yes	Yes	Yes	Yes
Industry fixed effect	Yes	Yes	Yes	Yes	Yes
Year fixed effect	Yes	Yes	Yes	Yes	Yes
Observations	12,254	12,234	8,398	12,234	11,452
Adjusted R-squared	0.513	0.913	0.908	0.913	0.907

Robust standard errors clustered by firm and year are in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 5: Net Buying Behavior-Based Measure of Overconfidence and TFP LP

This table presents the results when regressing TFP^{LP} on $NET\ BUYER$ along with control variables. All variables are defined in Table 1.

VARIABLE	Dependent Variable: TFP^{LP}				
	(1)	(2)	(3)	(4)	(5)
<i>NET BUYER</i>	0.147*** (0.012)	0.056*** (0.010)	0.028** (0.011)	0.056*** (0.010)	0.061*** (0.011)
<i>CEO AGE</i>	-0.006*** (0.001)	-0.002*** (0.000)	-0.003*** (0.001)	-0.002*** (0.000)	-0.002*** (0.000)
<i>CEO OWNERSHIP</i>	0.593*** (0.085)	0.299*** (0.082)	0.486*** (0.118)	0.300*** (0.082)	0.296*** (0.088)
<i>CEO FEMALE</i>	-0.074*** (0.023)	-0.018 (0.019)	0.015 (0.023)	-0.019 (0.019)	-0.020 (0.019)
<i>ROA</i>	0.923*** (0.070)	1.286*** (0.054)	1.143*** (0.064)	1.294*** (0.054)	1.223*** (0.054)
<i>TOBIN'S Q</i>	0.032*** (0.004)	0.018*** (0.003)	0.019*** (0.003)	0.018*** (0.003)	0.018*** (0.003)
<i>SIZE</i>	0.053*** (0.004)	0.087*** (0.007)	0.091*** (0.009)	0.089*** (0.007)	0.092*** (0.007)
<i>LEVERAGE</i>	0.194*** (0.027)	0.017 (0.024)	-0.042 (0.028)	0.016 (0.024)	0.008 (0.025)
<i>CASH</i>	0.765*** (0.057)	0.105*** (0.038)	0.104** (0.044)	0.104*** (0.038)	0.133*** (0.039)
<i>R&D</i>			0.057 (0.195)		
<i>R&D (missing values replaced by 0)</i>				0.121 (0.187)	
<i>R&D (missing values replaced by industry averages)</i>					0.139 (0.178)
Constant	0.754*** (0.050)	0.513*** (0.066)	0.527*** (0.086)	0.495*** (0.068)	0.467*** (0.071)
Firm fixed effect	No	Yes	Yes	Yes	Yes
Industry fixed effect	Yes	Yes	Yes	Yes	Yes
Year fixed effect	Yes	Yes	Yes	Yes	Yes
Observations	11,002	10,996	7,502	10,996	10,286
Adjusted R-squared	0.524	0.920	0.917	0.920	0.916

Robust standard errors clustered by firm and year are in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 6: CEO Change Analysis, Option-Based Measure of Overconfidence and TFP

This table presents the results for the subsample including firms that within the firms, there are switches from non-overconfident to overconfident CEOs and/or from overconfident to non-overconfident CEOs. The CEO overconfidence is measured based on *OVERCONFID67*. The table reports the results when regressing TFP^{LP} on *OVERCONFID67* along with control variables. All variables are defined in Table 1.

VARIABLE	Dependent Variable: TFP^{LP}				
	(1)	(2)	(3)	(4)	(5)
<i>OVERCONFID67</i>	0.033** (0.013)	0.055*** (0.015)	0.033*** (0.008)	0.045*** (0.007)	0.045*** (0.007)
<i>CEO AGE</i>	-0.001 (0.001)	-0.001 (0.001)	-0.001 (0.001)	-0.001* (0.001)	-0.001** (0.001)
<i>CEO OWNERSHIP</i>	0.261* (0.142)	0.528*** (0.146)	-0.064 (0.133)	0.144 (0.126)	0.161 (0.138)
<i>CEO FEMALE</i>	-0.108*** (0.026)	-0.103*** (0.023)	-0.030* (0.016)	-0.037** (0.016)	-0.034** (0.016)
<i>ROA</i>	1.171*** (0.104)	1.189*** (0.138)	0.893*** (0.091)	1.058*** (0.075)	1.054*** (0.078)
<i>TOBIN'S Q</i>	-0.003 (0.006)	-0.022*** (0.008)	0.020*** (0.005)	0.009** (0.004)	0.010** (0.005)
<i>SIZE</i>	0.063*** (0.005)	0.063*** (0.006)	0.043*** (0.012)	0.062*** (0.010)	0.067*** (0.010)
<i>LEVERAGE</i>	0.095** (0.038)	0.071 (0.045)	0.053 (0.041)	0.117*** (0.036)	0.113*** (0.037)
<i>CASH</i>	0.567*** (0.082)	0.360*** (0.097)	0.123* (0.070)	0.222*** (0.065)	0.234*** (0.067)
<i>R&D</i>		1.108*** (0.243)	-0.714* (0.385)		
<i>R&D (missing values replaced by 0)</i>				-0.432 (0.382)	
<i>R&D (missing values replaced by industry averages)</i>					-0.338 (0.326)
Constant	0.566*** (0.067)	0.632*** (0.082)	0.802*** (0.120)	0.600*** (0.094)	0.575*** (0.098)
Firm fixed effect	No	Yes	Yes	Yes	Yes
Industry fixed effect	Yes	Yes	Yes	Yes	Yes
Year fixed effect	Yes	Yes	Yes	Yes	Yes
Observations	4,883	3,372	3,370	4,879	4,623
Adjusted R-squared	0.495	0.480	0.903	0.891	0.891

Robust standard errors clustered by firm and year are in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 7: CEO Change Analysis, Net Buying Behavior-Based Measure of Overconfidence and TFP

This table presents the results for the subsample including firms that within the firms, there are switches from non-overconfident to overconfident CEOs and/or from overconfident to non-overconfident CEOs. The CEO overconfidence is measured based on *NET BUYER*. The table reports the results when regressing TFP^{LP} on *NET BUYER* along with control variables. All variables are defined in Table 1.

VARIABLE	Dependent Variable: TFP^{LP}				
	(1)	(2)	(3)	(4)	(5)
<i>NET BUYER</i>	0.059*** (0.019)	0.060*** (0.022)	0.032*** (0.010)	0.058*** (0.010)	0.058*** (0.011)
<i>CEO AGE</i>	-0.010*** (0.001)	-0.008*** (0.002)	-0.000 (0.001)	-0.002** (0.001)	-0.001* (0.001)
<i>CEO OWNERSHIP</i>	0.134 (0.138)	-0.046 (0.325)	0.220 (0.212)	0.106 (0.104)	0.077 (0.122)
<i>CEO FEMALE</i>	0.077* (0.042)	0.060 (0.045)	0.095* (0.050)	0.066 (0.040)	0.061 (0.040)
<i>ROA</i>	0.970*** (0.140)	0.979*** (0.211)	1.336*** (0.138)	1.375*** (0.109)	1.366*** (0.113)
<i>TOBIN'S Q</i>	0.009 (0.009)	-0.009 (0.012)	0.028*** (0.010)	0.022*** (0.008)	0.020** (0.008)
<i>SIZE</i>	0.061*** (0.009)	0.100*** (0.011)	0.068*** (0.019)	0.059*** (0.014)	0.065*** (0.015)
<i>LEVERAGE</i>	0.147*** (0.046)	0.102* (0.058)	0.133** (0.066)	0.141*** (0.047)	0.160*** (0.050)
<i>CASH</i>	0.097 (0.116)	-0.233 (0.156)	0.081 (0.099)	0.125 (0.080)	0.145* (0.086)
<i>R&D</i>		2.242*** (0.350)	-1.634** (0.705)		
<i>R&D (missing values replaced by 0)</i>				-1.636** (0.696)	
<i>R&D (missing values replaced by industry averages)</i>					-1.358** (0.548)
Constant	1.018*** (0.098)	0.599*** (0.131)	0.397** (0.170)	0.553*** (0.130)	0.511*** (0.138)
Firm fixed effect	No	Yes	Yes	Yes	Yes
Industry fixed effect	Yes	Yes	Yes	Yes	Yes
Year fixed effect	Yes	Yes	Yes	Yes	Yes
Observations	2,487	1,593	1,592	2,486	2,265
Adjusted R-squared	0.703	0.523	0.904	0.935	0.930

Robust standard errors clustered by firm and year are in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Appendix A: The Result in Condition 18

This Appendix is to present the process to obtain the result in Condition 18.

$$E \leq E_{max} = U(e > 0) - U(e = 0)$$

$$\begin{aligned}
&= \left(\frac{1}{2}(1 + \Delta)\bar{p} + \frac{1}{2}(1 - \Delta)\underline{p} \right) \frac{\bar{p}}{\bar{p} + \underline{p}} \bar{c} + \left(\frac{1}{2}(1 + \Delta)(1 - \bar{p}) + \frac{1}{2}(1 - \Delta)(1 - \underline{p}) \right) \underline{w} - \frac{1}{2} \bar{c} \\
&= \left(\frac{1}{2}(1 + \Delta)\bar{p} + (1 - \frac{1}{2}(1 + \Delta))\underline{p} \right) \frac{\bar{p}}{\bar{p} + \underline{p}} \bar{c} + \left(\frac{1}{2}(1 + \Delta)(1 - \bar{p}) + (1 - \frac{1}{2}(1 + \Delta))(1 - \underline{p}) \right) \underline{w} - \frac{1}{2} \bar{c} \\
&= \left(\frac{1}{2}(1 + \Delta)\bar{p} + \underline{p} - \frac{1}{2}(1 + \Delta)\underline{p} \right) \frac{\bar{p}}{\bar{p} + \underline{p}} \bar{c} + \left(\frac{1}{2}(1 + \Delta) - \frac{1}{2}(1 + \Delta)\bar{p} + 1 - \underline{p} - \frac{1}{2}(1 + \Delta) + \frac{1}{2}(1 + \Delta)\underline{p} \right) \underline{w} - \frac{1}{2} \bar{c} \\
&= \left(\frac{1}{2}(1 + \Delta)\bar{p} - \frac{1}{2}(1 + \Delta)\underline{p} \right) \left(\frac{\bar{p}}{\bar{p} + \underline{p}} \bar{c} - \underline{w} \right) + \underline{p} \frac{\bar{p}}{\bar{p} + \underline{p}} \bar{c} + \underline{w} - \underline{p}\underline{w} - \frac{1}{2} \bar{c} \\
&= \frac{1}{2}(1 + \Delta) (\bar{p} - \underline{p}) \left(\frac{\bar{p}}{\bar{p} + \underline{p}} \bar{c} - \underline{w} \right) + \underline{p} \left(\frac{\bar{p}}{\bar{p} + \underline{p}} \bar{c} - \underline{w} \right) + \underline{w} - \frac{1}{2} \bar{c} \\
&= \frac{1}{2} (\bar{p} - \underline{p}) \left(\frac{\bar{p}}{\bar{p} + \underline{p}} \bar{c} - \underline{w} \right) + \frac{1}{2} \Delta (\bar{p} - \underline{p}) \left(\frac{\bar{p}}{\bar{p} + \underline{p}} \bar{c} - \underline{w} \right) + \underline{p} \left(\frac{\bar{p}}{\bar{p} + \underline{p}} \bar{c} - \underline{w} \right) + \underline{w} - \frac{1}{2} \bar{c} \\
&= \left(\frac{1}{2} (\bar{p} - \underline{p}) + \underline{p} \right) \left(\frac{\bar{p}}{\bar{p} + \underline{p}} \bar{c} - \underline{w} \right) + \frac{1}{2} \Delta (\bar{p} - \underline{p}) \left(\frac{\bar{p}}{\bar{p} + \underline{p}} \bar{c} - \underline{w} \right) + \underline{w} - \frac{1}{2} \bar{c} \\
&= \left(\frac{1}{2} \bar{p} + \frac{1}{2} \underline{p} \right) \left(\frac{\bar{p}}{\bar{p} + \underline{p}} \bar{c} - \underline{w} \right) + \frac{1}{2} \Delta (\bar{p} - \underline{p}) \left(\frac{\bar{p}}{\bar{p} + \underline{p}} \bar{c} - \underline{w} \right) + \underline{w} - \frac{1}{2} \bar{c} \\
&= \frac{1}{2} (\bar{p} + \underline{p}) \left(\frac{\bar{p}}{\bar{p} + \underline{p}} \bar{c} - \underline{w} \right) + \frac{1}{2} \Delta (\bar{p} - \underline{p}) \left(\frac{\bar{p}}{\bar{p} + \underline{p}} \bar{c} - \underline{w} \right) + \underline{w} - \frac{1}{2} \bar{c} \\
&= \frac{1}{2} \bar{p} \bar{c} - \frac{1}{2} (\bar{p} + \underline{p}) \underline{w} + \frac{1}{2} \Delta (\bar{p} - \underline{p}) \left(\frac{\bar{p}}{\bar{p} + \underline{p}} \bar{c} - \underline{w} \right) + \underline{w} - \frac{1}{2} \bar{c} \\
&= \frac{1}{2} \bar{p} \bar{c} + \left(1 - \frac{1}{2} \bar{p} - \frac{1}{2} \underline{p} \right) \underline{w} + \frac{1}{2} \Delta (\bar{p} - \underline{p}) \left(\frac{\bar{p}}{\bar{p} + \underline{p}} \bar{c} - \underline{w} \right) - \frac{1}{2} \bar{c}
\end{aligned}$$

Appendix B: The Levinsohn and Petrin (LP) Estimation

The LP estimation of TFP using intermediate inputs (for example, materials) as the proxy requires two stages.

The first stage

Let assume the production function follows the Cobb-Douglas form which is:

$$\ln(Y_{it}) = \beta_0 + \beta_K \ln(K_{it}) + \beta_L \ln(L_{it}) + \beta_M \ln(M_{it}) + \theta_{it} + \mu_{it} \quad (1B)$$

In Equation 1B, the subscripts i and t denote firm i and year t , respectively. For simplicity, let denote $\ln(Y_{it})$, $\ln(K_{it})$, $\ln(L_{it})$, $\ln(M_{it})$ as y_{it} , k_{it} , l_{it} , and m_{it} , respectively. Then Equation 1B can be rewritten as:

$$y_{it} = \beta_0 + \beta_K k_{it} + \beta_L l_{it} + \beta_M m_{it} + \theta_{it} + \mu_{it} \quad (2B)$$

In Equation 2B, θ_{it} is the productivity and μ_{it} is the error term (Olley and Pakes 1996; Levinsohn and Petrin 2003). Levinsohn and Petrin (2003) use materials as the proxy and write materials as a demand function of the firm's capital and productivity:

$$m_{it} = m_{it}(k_{it}, \theta_{it}) \quad (3B)$$

Levinsohn and Petrin (2003) assume m_{it} is monotonically increasing in θ_{it} . The monotonicity assumption allows the materials' demand function can be inverted to obtain the function for θ_{it} :

$$\theta_{it} = \theta_{it}(k_{it}, m_{it}) \quad (4B)$$

Equation 2B then can be rewritten as:

$$y_{it} = \beta_L l_{it} + \varphi_{it}(k_{it}, m_{it}) + \mu_{it} \quad (5B)$$

Where:

$$\varphi_{it}(k_{it}, m_{it}) = \beta_0 + \beta_K k_{it} + \beta_M m_{it} + \theta_{it}(k_{it}, m_{it}) \quad (6B)$$

Levinsohn and Petrin (2003) noted that $\varphi_{it}(k_{it}, m_{it})$ can be approximated by a polynomial expansion in k_{it} and m_{it} . Petrin et al. (2004) use a third order polynomial approximation of $\varphi_{it}(k_{it}, m_{it})$, then Equation 6B can be rewritten as:

$$\varphi_{it}(k_{it}, m_{it}) = \delta_0 + \sum_{e=0}^3 \sum_{f=0}^{3-e} \delta_{ef} k_{it}^e m_{it}^f \quad (7B)$$

Plug (7B) into (5B), one obtains:

$$y_{it} = \delta_0 + \beta_L l_{it} + \sum_{e=0}^3 \sum_{f=0}^{3-e} \delta_{ef} k_{it}^e m_{it}^f + \mu_{it} \quad (8B)$$

One can use OLS to estimate (8B) and obtain the estimates of β_L and $\varphi_{it}(k_{it}, m_{it})$.

The second stage

After obtaining $\widehat{\beta}_L$ and $\widehat{\varphi}_{it}$ in the first stage, and for any candidate values β_K^* and β_M^* for β_K and β_M , respectively, then based on Equation 6B, Levinsohn and Petrin (2003) and Petrin et al. (2004) estimate $\widehat{\theta}_{it}$ using the following equation:

$$\widehat{\theta}_{it} = \widehat{\varphi}_{it} - \beta_K^* k_{it} - \beta_M^* m_{it} \quad (9B)$$

Levinsohn and Petrin (2003) also assume that the productivity term θ_{it} follows a first-order Markov process:

$$\theta_{it} = E[\theta_{it} | \theta_{it-1}] + \xi_{it} \quad (10B)$$

In Equation 10B, ξ_{it} is an innovation to productivity that might be correlated with l_{it} but is uncorrelated with k_{it} (Levinsohn and Petrin 2003; Petrin et al. 2004). Plugging (10B) into (2B) and one obtains:

$$y_{it} = \beta_0 + \beta_K k_{it} + \beta_L l_{it} + \beta_M m_{it} + E[\theta_{it} | \theta_{it-1}] + \xi_{it} + \mu_{it} \quad (11B)$$

The next step is to estimate the value of $E[\theta_{it} | \theta_{it-1}]$. To do so, Levinsohn and Petrin (2003) and Petrin et al. (2004) run the following regression:

$$\theta_{it} = \gamma_0 + \gamma_1 \theta_{it-1} + \gamma_2 \theta_{it-2}^2 + \gamma_3 \theta_{it-3}^3 + \epsilon_t \quad (12B)$$

After that, Levinsohn and Petrin (2003) and Petrin et al. (2004) obtain the predicted values of the Regression 12B and call the predicted values as $E[\theta_{it} | \widehat{\theta}_{it-1}]$. Then using Equation 11B, Levinsohn and Petrin (2003) and Petrin et al. (2004) estimate the residual $\xi_{it} + \mu_{it}$ as follows:

$$\widehat{\xi_{it} + \mu_{it}} = y_{it} - \widehat{\beta}_L l_{it} - \beta_K^* k_{it} - \beta_M^* m_{it} - E[\theta_{it} | \widehat{\theta}_{it-1}] \quad (13B)$$

To identify β_K , Levinsohn and Petrin (2003) make an important assumption that is the capital stock in period t , k_{it} , only depends on the investment in the previous period and does not depend on the productivity innovation in this period, meaning that:

$$E[\xi_{it} + \mu_{it} | k_{it}] = 0 \quad (14B)$$

Similarly, to identify β_M , Levinsohn and Petrin (2003) assumes that the materials choice in the last period, m_{it-1} , is unrelated to the productivity innovation in this period. This means:

$$E[\xi_{it} + \mu_{it} | m_{it-1}] = 0 \quad (15B)$$

Then Levinsohn and Petrin (2003) and Petrin et al. (2004) impose additional overidentification conditions:

$$E[\xi_{it} + \mu_{it} | l_{it-1}] = 0, E[\xi_{it} + \mu_{it} | m_{it-2}] = 0, \text{ and } E[\xi_{it} + \mu_{it} | k_{it-1}] = 0 \quad (16B)$$

Levinsohn and Petrin (2003) and Petrin et al. (2004) define $Z_{it} \equiv (k_{it}, m_{it-1}, l_{it-1}, m_{it-2}, k_{it-1})$. Finally, the estimates of β_K and β_M , $\widehat{\beta}_K$ and $\widehat{\beta}_M$, are the solutions to the following minimization problem:

$$\min_{(\beta_K, \beta_M)} \sum_h \left(\sum_t (\widehat{\xi_{it} + \mu_{it}}) Z_{hit} \right)^2 \quad (17B)$$

After obtaining $\widehat{\beta}_L$ in the first stage, $\widehat{\beta}_K$ and $\widehat{\beta}_M$ in the second stage, one can calculate the TFP in the log form using:

$$\widehat{\theta}_{it} = y_{it} - \widehat{\beta}_K k_{it} - \widehat{\beta}_L l_{it} - \widehat{\beta}_M m_{it} \quad (18B)$$

To facilitate the estimation, Petrin et al. (2004) develop the Stata command, `levpet`, to implement the procedure.