Evaluating an Investment Subsidy Policy through a Structural Econometric Model using Micro-data from Greece

Alexandros Fakos [a,*]

[a]Instituto Tecnológico Autónomo de México

February 15, 2016[†]
Current Version: `http://goo.gl/xh3sJw`

Abstract

This paper studies the effects of investment subsidies on firm's dynamic investment choices and aggregate investment. In particular, I focus on a subsidy program in which the Greek government subsidizes the acquisition of capital for selected firms. To analyze this program I develop a theoretical dynamic model of a firm's decision to invest in physical capital in the presence of the investment subsidy policy which effectively changes the price of capital for subsidized firms and the expectations about future capital prices for all firms. Using a novel micro-level dataset for the Greek food and beverages manufacturing sector that distinguishes subsidized from unsubsidized firms, I estimate the effect of the policy on firm investment behavior with two methodological approaches. A reduced-form firm investment demand equation is specified to estimate the average treatment effect of the policy on investment demand controlling for unobserved productivity. Exploiting variation in the intensity of the policy across locations and time a difrerences-in-differences investment demand regression is estimated. The results of this diff-in-diff analysis indicate that the policy has an economically significant effect on the investment behavior of unsubsidized firms (untreated population). The theoretical dynamic model is extended to a structural econometric model for estimating policy-invariant parameters by exploiting the firm's optimality conditions. I then use these estimated parameters to simulate alternative counterfactual policy scenarios. Consistent with the reduced-form evidence, the estimated structural model

*Email: `alexandrosfakos@gmail.com`. I thank my advisors Mark Roberts, Paul Grieco, and Stephen Yeaple for guidance, helpful discussions, and comments. I also want to thank Russell Cooper and Kenneth Judd for helpful suggestions. Seminar participants at The Pennsylvania State University provided helpful comments as well. I thank Penn State's Institute for CyberScience for technical support and computing time. I am grateful to the George and Victoria Karelias Foundation for financial support. The views expressed herein are those of the author and not necessarily those of the Hellenic Statistical Authority or the Greek Ministry of Development. All remaining errors are mine.

†`Version: 0.3.3`

predicts that subsidized firms invest more because of the policy while unsubsidized firms invest less because the policy increases their option value of waiting. I also find that subsidies are not randomly allocated among firms as the the probability of being subsidized increases with productivity.

JEL Classification Codes: D04, D22, C51, C63, H25, H32, L60
Keywords: Policy Evaluation, Subsidies, Investment, Capital Adjustment Cost, Dynamic Inputs, Productivity, Dynamic Treatment Effects, Industrial Policy

1 Introduction

Firm investment behavior determines two of the most important and volatile components of aggregate activity: aggregate investment directly and aggregate total factor productivity indirectly through the allocation of capital across heterogeneous producers. In a recent study, Khan and Thomas (2008) argue that, in the presence of persistent firm heterogeneity, explaining cyclical fluctuations in aggregate investment requires an understanding of individual firm behavior and its determinants, namely capital adjustment costs and the nature of idiosyncratic productivity shocks. Capital adjustment costs, investment irreversibility, and the lag between the investment decision and the operationalization of the new capital make capital a dynamic input thus, the investment decision becomes crucially dependent on expectations about the future. Quantitative contributions by Restuccia and Rogerson (2008), Hsieh and Klenow (2009), and Midrigan and Xu (2014) document the substantial effects of capital misallocation on aggregate productivity while empirical investigations by Cooper and Haltiwanger (2006), Fuentes, Gilchrist and Rysman (2006), Bloom (2009), and Asker, Collard-Wexler and De Loecker (2014) demonstrate the importance of capital adjustment costs and investment irreversibility. Due to the key role of investment for the business cycle and additional important economic factors such as externalities and other market failures, policy makers and researchers seek to design policies to stimulate investment and evaluate their effectiveness.

In this paper I develop and estimate a dynamic model of firm investment choice using a novel dataset for Greek manufacturing plants in the food and beverages industry spanning over the period 1999-2009. I use the model to quantify the effect of an investment subsidy policy on firm's investment behavior controlling for unobserved heterogeneity. This policy is active during the whole sample period and is implemented mainly through cash investment subsidies; that is, the government bears a fraction of the cost of the investment projects undertaken by the subsidized firms as long as they complete the projects within a pre-specified time period. The investment subsidy program has two main goals: improve the economic conditions in the periphery of the country and give incentives to firms in the manufacturing and tourism sector to acquire state-of-the-art physical capital to become more competitive and produce high-quality products. The policy effectively changes the acquisition price of physical capital for subsidized firms and the expectations of all firms, subsidized and not, about the future price of capital. Since the subsidy program is already operational during the whole sample period, I need to control for its influence on observed firm behavior in order to recover the structural (policy-invariant) parameters of the model. These parameters allow for performing counterfactual experiments of interest such as the simulation of the aggregate impacts that a termination of the policy might exert over sectoral aggregate investment.

As part of my model I estimate the dynamic decision rule for the firms' optimal investment choice where this rule depends on expected future profits, the sunk cost of investment, and both the current price of capital and firm's expectation about the future price of capital. The latter directly depends on the subsidy policy. I analyze three pathways linking investment, subsidy rates, and productivity. First, the return to investment increases with the firm's underlying productivity, which leads highly productive firms to invest more. Second, subsidies increase subsidized firms' propensity to invest by directly decreasing their current cost of investment. Third, the fact that the expected future price of capital is higher than the current price for subsidized firms and lower than the current price for unsubsidized ones, reinforces the subsidized firms' incentives to exploit the subsidy rate investing intensively and induces unsubsidized firms to invest less than they would in the absence of the policy.

My estimation strategy relies on a dynamic stochastic optimization problem of an individual firm facing uncertainty in the future streams of profits and future prices of the capital good affected by the subsidy policy. Using firm's optimality conditions, which are calculated numerically, I estimate the parameters characterizing capital adjustment cost and the policy allocation rule with maximum likelihood. My results indicate that there is substantial heterogeneity in productivity across firms and failing to control for it can lead to incorrect inference about the effectiveness of the policy. I also find that subsidies are not randomly allocated among firms as the probability of being subsidized increases with productivity. Using the estimated parameters of the model I perform preliminary counterfactual simulations which suggest that the subsidy policy decreased sectoral aggregate investment by 11 percent. This to some extent surprising result depends on the fact that the subsidy policy induces highly productive, unsubsidized firms with high probability of being subsidized in the future to invest less. The increase in investment by subsidized firms due to the policy is not enough to compensate for the negative effect of the policy on unsubsidized firms.

Two insights arise from the analysis: The first is that evaluating the effect of the policy using unsubsidized/untreated firms as a control group in a regression design can be misleading and overestimate the effect of the policy because the policy can reduce the investment of unsubsidized firms. The effect of the policy on the untreated population arises even in a partial equilibrium framework, where factor prices remain constant, because of the dynamic nature of investment. The second insight is the importance of heterogeneity in calculating counterfactual scenarios. High-productivity firms are strongly affected by the policy while low-productivity firms are not which implies that in order to estimate the effect of the policy on aggregate investment researchers need to know the distribution of unobserved productivity.

The paper is organized as follows. The next section provides an overview of the relevant literature on firm investment and policy evaluation while section 3 briefly describes the Greek subsidy policy. Section 4 develops the theoretical model of the firm's dynamic decision to invest in physical capital, and section 5 describes the data used to carry out the empirical analysis. Section 6 describes the methodology to estimate the firm's profit function and unobserved productivity and presents the estimates. Building on the theoretical mode and utilizing the productivity estimates, section 7 specifies a reduced-form investment demand equation to estimate the average treatment effect of the policy on investment demand controlling for unobserved productivity and presents evidence of the effect of the subsidy policy on unsubsidized/untreated firms. Section 8 extends the theoretical dynamic model to an econometric model for estimating policy-invariant parameters. These estimated parameters are then used to simulate a counterfactual policy scenario where the subsidy policy is not in effect. The last section concludes with a summary and a discussion about the future directions of this research project.

2 Literature review

This paper is related to the literature of empirical studies developing fully specified dynamic models to analyze firms investment behavior using firm- or plant-level data. It is also related to a body of literature utilizing either aggregate or micro-data to evaluate economic policies aimed at stimulating investment in durable assets by firms and households. In this section, I provide a brief description of these two strands of literature and and their connection to my work.

Economists have always been interested in the forces behind the determination of investment and fixed capital and it has been recognized at least since Fisher (1907) that the decision to acquire fixed capital is fundamentally different than the decision to purchase perishable goods due to capital's durable character. Dixit and Pindyck (1994) lay out the modern theory of firm investment in which the sunk cost of investment, in combination with the uncertainty over future profits, the time it takes for capital to become operational, and capital's depreciation rate, determine firm's optimal investment choice. Researchers carrying out empirical studies on firm investment have to decide at first which variables to include in a model describing firm's decision to acquire physical capital. To this end, Modigliani and Miller (1958) provide some guidance by showing that under the assumption of efficient financial markets a firm's financial structure is irrelevant for its investment choice, thus the econometrician can ignore financial variables in the analysis. The assumptions of the Modigliani-Miller theorem are assumed to hold in a large part of the empirical literature studying firm's investment decisions and are assumed in my analysis as well[1].

Numerous industry case studies such as Peck (1974) and investigations of disaggregated plant-level datasets such as Doms and Dunne (1998) indicate that investment adjustment costs are needed to rationalize observed firms' investment behavior[2]. In fact, estimating the parameters determining adjustment costs from observed firm investment behavior has been the focus of several studies. Hayashi (1982) and Gilchrist and Himmelberg (1995) rely on the Q theory of investment and panel data methods while Caballero and Engel

[1]For empirical studies attempting to test the validity of these assumptions see Hubbard, Kashyap and Whited (1995), Fazzari, Hubbard and Petersen (2000) and the references therein.

[2]For more references to such studies see Cooper and Haltiwanger (2006).

(1999) estimate a dynamic model with non-convex capital adjustment costs via maximum likelihood from aggregate data. Cooper and Haltiwanger (2006) and Asker, Collard-Wexler and De Loecker (2014) also develop fully specified dynamic models and indirect inference methods to to recover flexible capital adjustment cost functions using firm-level data.

Knowledge of the underlying capital adjustment costs is necessary to predict firms' investment response to exogenous shocks such as aggregate productivity shocks, a tax or a subsidy policy. For instance, Bloom (2009) shows that during periods of high uncertainty, firms facing capital adjustment costs can be extremely insensitive to price changes rendering monetary or fiscal policies potentially ineffective in stimulating investment. Moreover, recent contributions by Restuccia and Rogerson (2008), Hsieh and Klenow (2009), Midrigan and Xu (2014) construct measures of capital misallocation using the cross-sectional distribution of capital and productivity across firms and Asker, Collard-Wexler and De Loecker (2014) examine the quantitative challenges of measuring capital misallocation and stress the importance of accurately measuring capital adjustment costs. My work builds on this literature by developing a dynamic model to recover firms' capital adjustment cost in order to evaluate the effects of a Greek investment subsidy policy on firm behavior. Since my paper uses data to evaluate the effect of a realized government policy, it is also related to the literature on policy evaluation.

For many reasons that are beyond the scope of this paper, stimulating investment is usually part of policy makers' agenda. Investment policies typically target firm investment in physical capital and R&D as well as household investment in fixed assets through either tax incentives or direct investment subsidies. Two bodies of literature have been developed to study the effects of such policies. One strand utilizes quasi-experimental data spanning periods during which a government policy is in effect to measure what the outcome would have been in the absence of the policy. For example, Criscuolo, Martin, Overman and Van

Reenen (2012) study the effect of the Regional Selective Assistance program in the UK on regional employment and investment. House and Shapiro (2008) use aggregate data to recover the properties of supply and demand in the market for investment goods and analyze the response of capital prices to a temporary bonus depreciation tax incentive in the USA. Yagan (2014) exploits the quasi-experimental nature of the dividend tax cut of the Jobs and Growth Tax Relief Reconciliation Act to quantify its effect on real investment. Mian and Sufi (2012) use aggregate data to study the effect of the 'Cash for Clunkers' program in the USA on automobile purchases. This strand can be referred to as ex-post policy evaluation literature. Another body of literature performs what can be defined as ex-ante policy evaluation in which a specific model is constructed and applied to data spanning policy-free periods to estimate structural (policy-invariant) parameters governing agents' behavior. These estimates can be used to predict the impact of various counterfactual policy scenarios. To this extent, Adda and Cooper (2000) study how a temporary automobile scraping subsidies in France affected the car age distribution and aggregate car purchases in the short and the long run.

My analysis bears similarities with both strands of the policy evaluation literature. The methodology and the questions are more closely related to the ex-ante policy evaluation literature as my goal is to assess the effect of variations of the policy on aggregate outcomes. In addition, I face the challenges that the ex-post policy evaluation approach presents. The policy in question is in effect during the whole sample period, which implies that observed investment behavior is affected by the policy in place. Furthermore, the subsidized firms are not a random subset of the heterogeneous firm population, hence I need to account for selection. Lastly, this paper provides evidence of the effect of economic policies on the untreated population like Angelucci and Giorgi (2009) who find evidence that a Mexican welfare program of cash transfers, Progresa, had an effect on the untreated households' consumption.

3 The Greek investment subsidy program

The Greek government has been implementing an investment incentive program for firms in the manufacturing and tourism sector since 1981. The program has two main goals: improve the economic conditions in the periphery of the country and give incentives to firms to acquire state-of-the-art physical capital to become more competitive and produce high-quality products[3]. The dominant policy instrument used to achieve these goals is cash investment subsidies; that is, the government bears a fraction of the cost of an investment project the subsidized firm undertakes[4]. Moreover, the program is regional in character since firms located in specific regions face higher subsidy rates and have higher probability of being subsidized. The program has always had a sizable budget but its scope expanded in the second half of the 2000s. Figure 3 depicts the total amount of cash subsidies granted by the program and the median subsidy rate granted. Notice that the cumulative amount of cash subsidies granted in 2006 is roughly a billion Euros which is approximately 2 percent of the gross fixed capital formation in the whole economy. The median subsidy rate also varies across time fluctuating between 30 and 45 percent with an upward trend from 2005 onwards.

The budget and implementation details of the subsidy program is set at the national level by laws voted in the parliament and revised periodically. The laws specifying the institutional details of the subsidy program do not lapse and remain in effect until the next revision. During the period 1998 - 2009 two laws came into effect (specifically in 1998 and

[3] For several years the program also heavily subsidized the construction of solar and wind power generating farms.

[4] Tax credits and subsidization of interest payments are two other instruments but approximately 95 percent of the subsidized firms during the period 1998-2010 covered in the analysis received aid in the form of a cash subsidy.

2004) establishing that at any point in time firms could apply for a subsidy by submitting an investment project to the authority implementing the subsidy program. Subsidized firms have a specified amount of time to complete the project. Upon the deadline for the project completion and any extension the firm might have obtained, the firm receives a cash transfer equivalent to the subsidy rate times the investment expenditure[5]. Firms do not receive a cash transfer for any uncompleted part of the investment project.

4 A dynamic model of investment

The theoretical model presented in this section bears several similarities to the models of investment developed by Cooper and Haltiwanger (2006) and Asker, Collard-Wexler and De Loecker (2014). There are two key different assumptions in my model. The first is the addition of one extra state variable, the price of investment, which varies across firms, follows an exogenous Markov process and encompasses the government subsidy policy. The second is that investment is assumed to be completely irreversible. I abstract from the decision of the firm to enter or exit production and focus on the decision to invest in physical capital. Firms are heterogeneous in their total factor productivity, the price of investment and their location. Location affects directly firms' expectations about future prices and hence, along with productivity and current price of the capital good, affects their incentive to invest.

4.1 Static decisions

I begin by describing the production function for a representative firm j and the demand for its output which jointly determine its variable profits; the demand and marginal cost specification is as in Aw, Roberts and Xu (2011). In what follows $t = 1, \ldots, T$ is an

[5]Subsidized firms usually receive half of the cash subsidy after half of the investment project is realized.

indicator of time and the j subscript is omitted to simplify notation. Firm's marginal cost depends on the current level of capital, productivity and variable input prices and is specified as[6]

$$\ln c_t = \ln c(k_t, w_t) - \omega_t = \beta_0 + \beta_k \ln k_t + \beta_w \ln w - \omega_t \tag{1}$$

where k is physical capital stock, ω is productivity and w is a vector of the prevailing variable input prices. Demand for firm's output q_t takes the Dixit-Stiglitz form and is given by

$$q_t = Q\left(\frac{p_t}{P}\right)^{\eta} = \frac{I}{P}\left(\frac{p_t}{P}\right)^{\eta} = \Phi \cdot (p_t)^{\eta},\ \eta < -1$$

where Q is the industry output, P is the industry price index, p is the firm price of output, I is the market size and η is the demand elasticity. Firms choose the output price to maximize variable profits $(p_t - c_t)q_t$. Under the optimal pricing policy

$$p_t = \frac{\eta}{\eta + 1} c_t \tag{2}$$

variable profits are

$$\pi_t = -\frac{1}{\eta}\left(\frac{\eta}{\eta+1}\right)^{\eta+1} \Phi \left(e^{\beta_k \ln k_t + \beta_w \ln w - \omega_t}\right)^{\eta+1} \tag{3}$$

$$\pi(\omega_{it}, k_t) = Ae^{-(\eta+1)\omega_t} k_t^{\beta_k(\eta+1)} \tag{4}$$

$$\ln \pi_t = \ln(A) + (\eta + 1)(\beta_k \ln k_t - \omega_t) \tag{5}$$

[6]For the derivation of this cost function from a production function see appendix B.1.

with A being a term combining all the time effects that do not change across firms. Equation (4) shows that variable profits can provide information on firm's productivity level ω_t and illustrates the properties of the variable profit function. The demand parameter η and the supply parameter β_k jointly establish the curvature of the variable profit function with respect to capital. Its curvature is one of the driving forces behind firms' investment choice, which is the main focus of this paper. Heterogeneity in variable profits across firms arises from productivity ω and size[7] k. Firm size is the result of past investment choices and productivity captures all the other sources of variable profit heterogeneity such as managerial ability, product quality, and cost differences. Productivity is modeled as an exogenous stochastic process defined in the following section.

4.2 Transition of productivity ω

Productivity follows a first order Markov stochastic process and is independent of firm's choices. The assumption that firms cannot affect the evolution of productivity is a simplifying abstraction. Several recent papers focus on quantifying the linkages between R&D and productivity[8] but I ignore such linkages partly due to the absence of a specific measure of R&D in my dataset and partly to keep the model tractable and concentrate the analysis on the effect of the price of capital on investment choices. The Markov process is specified as:

$$\omega_{t+1} = E\left(\omega_{t+1}|\omega_t\right) + \xi_t = g(\omega_t) + \xi_t \tag{6}$$

where ξ is a mean zero *iid* random variable defining the Markov transition function $F_\omega(\omega_{t+1}|\omega_t)$ along with g. For each $\omega_t \in \mathbf{R}$, $F_\omega(\omega_{t+1}|\omega_t)$ is a probability distribution

[7] I use the terms 'size' and 'capital stock' interchangeably.

[8] See for example Aw, Roberts and Xu (2011), Doraszelski and Jaumandreu (2013).

with mean $g(\omega_t)$. The serially correlated productivity process generates a correlation between current productivity and future variable profits making current productivity one of the main determinants of firm's investment choice.

4.3 Dynamic decision to invest

In this section I present a model of the firm's decision to invest in physical capital which is closely related to the one developed in Asker, Collard-Wexler and De Loecker (2014) and in Fuentes, Gilchrist and Rysman (2006). The key assumption in my model is that investment is irreversible and newly purchased capital becomes operational one period after it is acquired rendering capital a dynamic input. Firms face different prices p_k of the capital good k due to the presence of an investment subsidy policy. Each firm arrives at the beginning of period t with a capital stock k_t, observes the realization of productivity ω_t, the price of capital p_{k_t} and an *iid* shock to the investment cost ϵ_t and, given this information, decides the optimal output price characterized in (2) and the level of investment in physical capital.

The state vector for each firm is $s_t = (\omega_t, p_{k_t}, k_t, l, \epsilon_t)$ where l is the time-invariant location of the firm. Location only affects expectations about future values of p_k because the subsidy policy can vary by location. Physical capital k is a quasi-fixed input, depreciates at rate δ and evolves in a deterministic fashion. More specifically, a firm investing quantity i_t at time t arrives at $t+1$ with capital

$$k_{t+1} = i_t + (1-\delta)k_t \tag{7}$$

thus, choosing the quantity of investment today i_t is equivalent to choosing next period capital stock k_{t+1} and the terms 'choice of investment' and 'choice of next period's capital' can be used interchangeably. This time-to-build characteristic of capital accumulation along with the *iid* property of the shock to the investment cost ϵ imply that today's investment has no effect on current variable profits $\pi(\omega_t, k_t)$, therefore output price and investment are decided independently.

The payoff of the firm at time t consists of two components: variable profits $\pi(\omega_t, k_t)$ and the cost of investing $[p_{k_t} i_t \mathbf{1}\{i_t > 0\} + C(k_t, k_{t+1})]\, \epsilon_t,\ \epsilon_t \sim F_\epsilon;\ \text{Prob}(\epsilon > 0) = 1$. The cost of investment is equal to the investment expenditure $p_{k_{it}} i_t$ plus a positive capital adjustment cost $C(k_t, k_{t+1})$, both multiplied by a factor ϵ_t. The shock ϵ is independent across firms and time and also independent of the other state variables. It represents idiosyncratic shocks to the cost of investment and it affects the optimal level of investment.

Note that if gross investment i is negative, the investment cost for the firm is $-C(k, k')$, which implies that selling existing capital stock not only generates zero revenue, but also requires the firm to incur an adjustment cost to dispose of it. In other words, any investment expenditure is a sunk cost. The firm's current payoff function is

$$U(s_t, k_{t+1}) = \pi(\omega_t, k_t) - [p_{k_t}(k_{t+1} - (1-\delta)k_t)\mathbf{1}\{i_t > 0\} - C(k_t, k_{t+1})]\, \epsilon_t \tag{8}$$

4.3.1 Transition of the Price of Capital

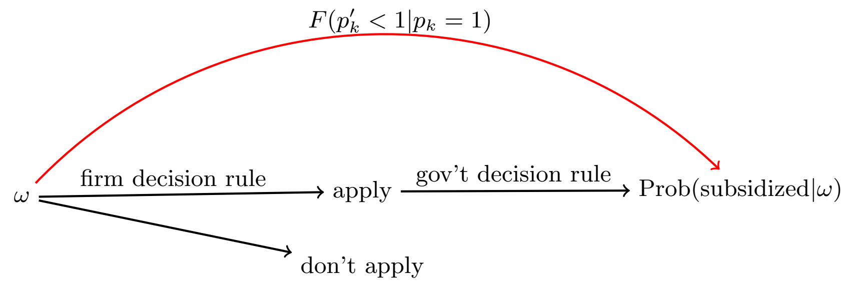
Before defining the firm's maximization problem it is crucial to characterize the stochastic process governing the evolution of the price of capital p_k. The presence of the price of capital in the state vector is a distinctive feature of the model, which is specifically designed to study the linkage between p_k and firms' investment decisions. The subsidy policy directly affects p_k because an investment subsidy is a proportional discount on the investment expenditure.

The price of capital p_k is normalized such that $p_k \in (0, 1]$, with an unsubsidized firm facing $p_k = 1$ and a subsidized firm facing price $p_k < 1$. For example, a firm receiving a 30 percent subsidy rate can purchase each unit of physical capital for price $p_k = .7$. The random variable p_k evolves according to a Markov process with transition function $F_{p_k}(p_{k_{t+1}}|p_{k_t}, l, \omega)$. The investment subsidy policy is completely characterized by F_{p_k} thus, in the context of the model, the transition function *is* the subsidy policy. Note that F_{p_k} depends on location because the subsidy policy analyzed here is location-dependent. Moreover, the probability distribution over next period's price $p_{k_{t+1}}$ depends on the current period's price p_{k_t} because firms that are subsidized today have a different probability to be subsidized tomorrow than firms that are unsubsidized today.

The probability distribution of $p_{k_{t+1}}$ also depends on the current productivity ω_t to account for both the self-selection of firms into the policy and a productivity-dependent government allocation rule. More specifically, the probability of receiving a subsidy in the following period $F(p_{k_{t+1}}|p_{k_t} = 1, l, \omega_t)$ is a reduced form of the subsidy allocation mechanism which involves the decision of firms to apply for a subsidy and the government rule for selecting which applications will receive a grant (see figure 1). The reason why I do not include the application process into the firm's problem is that I do not need to separate the self-selection of firms into the policy from the government allocation rule in order to perform the counterfactual experiment of interest which is the behavior of firms in the absence of the policy.

To summarize, at the beginning of each period t, given the location of the firm l, and the previous year's price $p_{k_{t-1}}$ and productivity ω_{t-1}, the current price p_{k_t} is drawn from the probability distribution $F_{p_k}(p_{k_{t+1}}|p_{k_t}, l, \omega_t)$.

Figure 1: Probability of receiving a subsidy



4.3.2 Firm's dynamic problem

In each period t, given the realization of the state s_t, each firm chooses an optimal investment policy $k_{t+1}(s_t)$ to maximize the expected discounted stream of payoffs

$$E\sum_{v=0}^{\infty}\beta^v U(s_{t+v}, k_{t+v+1})$$

where $\beta < 1$ is the discount factor and the expectation E is consistent with the Markov transition function $F_\omega(\omega_{t+1}|\omega_t)F_{p_k}(p_{k_{t+1}}|p_{k_t}, l, \omega_t)F_\epsilon(\epsilon_{t+1})$. If U is bounded, this dynamic programming problem has a recursive representation through a Bellman equation. Recall that investment is irreversible or, equivalently, investment is a sunk cost. The one year time-to-build, durability and sunk cost features of the capital accumulation process make investment a dynamic input, thus expectations about future states become pivotal in the investment decision. This investment irreversibility assumption is quite strong but useful to simplify the model in the presence of the subsidy policy. The main issue with relaxing this assumption is that if investment is reversible subsidized firms that can purchase capital at price $p_k < 1$ can sell the same capital next period at market price $p_k = 1$ realizing unbounded returns. In such economic environment, the prevailing price of capital in the market is directly affected by the subsidy rates implying that the subsidy policy has substantial general equilibrium effects that would be impossible to ignore in any counterfactual exercise. In reality, there are institutional constraints that prevent arbitrage opportunities in the market for physical capital. Such constraints could be imposed within the model by prohibiting firms to sell capital at a higher price than the purchase price. However, such a formulation[9] requires the inclusion of the history of all past capital purchases and purchase prices p_k in the state vector making the dynamic programming problem intractable.

Notice that, since investment is irreversible, firms never find it optimal to undertake negative investment and the firm's dynamic programming problem can be represented by the Bellman equation

$$\begin{aligned} V(\omega, p_k, k, l, \epsilon) = \quad & \max_{k' \geq (1-\delta)k} \pi(\omega, k) - \left[ip_k + C(k, k')\right] \epsilon + \\ & + \beta \int_{\omega', p_k', \epsilon'} V(\omega', p_k', k', l, \epsilon') dF(\omega', p_k' | \omega, p_k) dF_\epsilon(\epsilon') \\ & F(\omega', p_k' | s) = F_\omega(\omega' | \omega) F_{p_k}(p_k' | p_k, l, \omega) \\ & k' = i + (1-\delta)k, \; C(k, k') > 0 \end{aligned} \tag{9}$$

Because ϵ is serially independent, the current value of ϵ affects firm's decision to invest only through current cost of investment and has no effect on the firm's expectations about the future values of ϵ. This formulation of the dynamic problem implicitly assumes that there is no firm exit. This assumption is mainly due to data limitations because when a firm drops out of the sample it could simply be because it has less than 10 employees[10] and thus, it is not possible to know whether the firm exited the market or just downsized.

To summarize, firms are heterogeneous in their productivity, their capital stock, the price of capital they face, their location and an idiosyncratic cost to investment. The

[9] A reformulation of the state space such that all the capital vintages along with their purchase prices can be summarized by a sufficient statistic might be possible. Also redefining the firm's problem in order to capture more institutional details is part of ongoing research.

[10] Firms with less than 10 employees are not surveyed.

evolution of the price of capital, productivity and the idiosyncratic shock to investment cost is exogenous. Investment is irreversible, newly purchased capital becomes operational one year after its purchase and adjusting the capital stock is costly. Productivity and capital stock jointly determine the firm's short-run variable profits. Finally, the processes governing productivity, capital price, and idiosyncratic shock to investment in combination with capital price and adjustment cost drive firm's investment behavior. The timing of firm's decisions and arrival of information is depicted in figure 2.

Figure 2: Timing of decisions

t

- Arrive with capital k_t
- Observe the state s_t which includes productivity ω_t, the price of capital p_{k_t} and the idiosyncratic shock to the cost of investment ϵ_t.

- Decide the price of output to maximize current profits. Realized profits are $\pi(\omega_t, k_t)$
- Choose investment level $i_t \geq 0$ to maximize the future stream of profits.

$t+1$

- New capital becomes operational $k_{t+1} = (1-\delta)k_t + i_t$

4.4 Implications of the model

The theoretical model developed in the previous section has several implications which will be used in analyzing the data. Notice that both firm's choice variables, investment, and output price, depend on capital stock and productivity. The optimal investment is a function $i = i(\omega, p_k, k, l, \epsilon; F_\omega, F_{p_k}, F_\epsilon)$ while optimal output price is a function $p = p(\omega, k)$. The fact that the output pricing decision does not depend on the dynamic parameters and the subsidy policy depends on the time-to-build assumption which implies that investment decisions do not have any impact on current variable profits. This feature of the model will be exploited in the estimation of the structural econometric model and the reduced form analysis of section 7.

The theoretical model also implies that, typically, investment decisions of all firms, both subsidized and unsubsidized ones, are affected by changes in the subsidy policy F_{p_k}. Based on the real-options arguments of Dixit and Pindyck (1994), an unsubsidized firm facing $p_k = 1$ may make different investment decisions depending on the policy environment F_{p_k}. More specifically, an unsubsidized firm operating in an environment where there is a subsidy policy ($\text{Prob}(p_k' < 1|p_k = 1) > 0$) may invest less than it would in an environment where there is no subsidy policy ($\text{Prob}(p_k' < 1|p_k = 1) = 0$). This is because this firm has incentives to invest less today in the hope of having the option of acquiring capital at a lower price tomorrow. In my analysis I quantify this option value of waiting and this mechanism plays a crucial role in the counterfactual simulations.

Moreover, recall that a firm's location affects exclusively its investment decision indirectly through $F_{p_k}(p_k'|p_k, \omega, l)$. In the context of the model, the only reason why two firms whose state is identical except for location make different investment decisions is the fact that the subsidy policy varies across locations. I use this feature of the model to exploit the variation in the intensity of the subsidy policy across locations to explore the quantitative importance of the policy on the decisions of unsubsidized firms in section 7.

5 Data

The model developed in section 4 is used to analyze the effects of an investment subsidy program implemented by the Greek government on aggregate investment and productivity in the Greek food and beverage manufacturing sector. The micro-level data used in this paper come from three sources: the Hellenic Statistical Authority, the Greek Ministry of Development and a private firm named ICAP. Data on input and output flows come from the Annual Survey of Manufactures conducted by the Hellenic Statistical Authority at the plant level covering all manufacturing firms with at least 10 employees over the years 1999-2010. Variables include gross investment, total labor cost, number of employees, revenue, and expenditures on materials and electricity. The Greek Ministry of Development provided administrative micro data on subsidy grants at the firm level, namely the size of the subsidized expenditure, the cash transfer and the application and decision dates of each subsidy grant. The application date of unsuccessful applications is also included in the data but the rejection date is not included. Data on the book value of physical capital and the accumulated depreciation from firm's financial statements collected by ICAP are used in combination with investment flows from the Annual Survey of Manufactures to create the stock of physical capital at the firm level[11]. For the purpose of the empirical analysis I assume that any decision is taken at the firm level and the unit of observation is the firm. 80 percent of the firms owned a single plant and 11 percent owned at most two plants during the sample years while the rest of the firms owned at least 3 plants at some point during the sample period.

The dataset is an unbalanced panel of 430 firms accounting for 38 percent of the total sales of the sector in 2002[12]. There are many reasons why the panel is unbalanced: firms may stop being surveyed because their employment drops below the threshold, they exit from the market, or fail to report the required information for one or more years. Unfortunately, it is not possible to know which of these reasons caused the disappearance of a firm from the panel and this is why I do not attempt to explicitly model firm entry and exit. Food and beverages manufacturing is the largest sector in Greek manufacturing, both in terms of employment and value added, and received investment subsidies commensurately to its size. I exclude year 2010 from the analysis to avoid the large aggregate effect of the fiscal crisis in Greece on manufacturing firms. The model developed in this paper is a description of firms' behavior at a stationary equilibrium therefore it is not suitable for

[11] For more details on the construction of the variables and the estimation sample see appendix A.
[12] Based on data from the 2002 business registry.

the analysis of firms' behavior during episodes of extreme aggregate fluctuations[13].

Table 1 reports summary statistics of the input and output variables. The median firm employs 46 employees and the size (measured either in terms of capital stock or number of employees) distribution of firms is highly skewed towards small firms as demonstrated by the fact that the mean is two to three times the median for both capital stock and employment. Even if the correlation between capital stock and profits is positive and high (.75), it is not sufficiently close to 1 suggesting that allowing for some form of firm heterogeneity in necessary to explain the variation in profits. The theoretical model laid out in section 4 introduces firm heterogeneity through a serially correlated productivity process ω_t. Gross investment i has a probability mass at point zero as corroborated by the 6.3 percent observations with zero investment. This feature of the data is rationalized within the model by the sunk nature of investment in combination with the unobserved productivity process ω: large firms with a low productivity shock would like to downsize but, since the price of their already installed physical capital generates zero revenue if sold, their optimal choice is inaction.

5.1 Data patterns

This section provides a description of the subsidy policy data. The scope of the policy is economy wide but I present data only for the food and beverage manufacturing sector for which I perform the analysis. An observation in the subsidy data consists of the year the decision to subsidize a firm is made, the subsidy rate and the size of the subsidized project. In practice, each subsidized firm has a pre-specified number of years after the subsidy decision date to complete the subsidized project[14]. Upon the deadline for the

[13] Besides, a careful analysis of aggregate fluctuations at the firm level would require more than one year of data from the recession period and is left for future research.

[14] Extensions are possible but uncertain.

project completion and any extension the firm might have obtained, the firm receives a cash transfer equivalent to the subsidy rate times the investment expenditure[15]. Summary statistics of project size, subsidy rate, and decision year are presented in table 2. Notice that the subsidy rate is substantial, averaging at 40 percent, with values ranging from 5 to as high as 55 percent and more than half of the subsidy decisions are made during the period 2005-2008. The reason most of the subsidies are allocated from 2005 onwards is that the intensity of the policy changed in 2004. An overhaul of the subsidy program was voted in the Greek parliament and both the budget and the subsidy rates increased substantially. The median subsidized project size is 1.5 million Euros, which is roughly seven times higher than the median investment in the population. Even accounting for the fact that the implementation of subsidized projects can take more than one year, the relative size of the subsidized projects indicates that the policy could have an economically significant impact on individual investment behavior. In fact, the sum of all funded projects is 378 million Euros which is more that 10 percent of the aggregate investment by the firms in the sample, indicating that the subsidy policy has a budget that is sufficient to generate aggregate implications. Table 3 reports aggregate quantities for investment expenditure and subsidy funding. As shown, the observations corresponding to the period within 3 years of the reception of a subsidy, account for 11 percent of total aggregate investment in the sample even though these are only 7 percent of the sample observations. Once again, this s an indication of a possible aggregate impact of the policy.

To gain an understanding of the possible effect of the policy on firm's investment behavior, table 4 breaks down investment patterns for firms that received a grant at any point during the sample. Firm-year observations are categorized according to the time elapsed since the latest subsidy was granted to each firm. Observe that the median investment

[15]The firm does not receive a cash transfer for any uncompleted parts of the investment project.

level jumps after a subsidy is granted and goes back to the pre-subsidy level 3 years after the subsidy was granted. The pattern is similar for the median investment rate even though it is not as stark. While these patterns indicate that firms' investment behavior is probably influenced by the policy, it is not clear whether firms simply inter-temporally reallocate investment in the presence of the policy or the policy actually induces firms to invest substantially more.

Table 5 shows that there are differences between firms that are never subsidized and firms that are subsidized at some point even during the years prior to the reception of the subsidy grant. Subsidized firms invest twice as much as the unsubsidized ones, exhibit higher growth rates and are larger. These numbers indicate that subsidies are not allocated randomly across the population and my analysis should account for that. Moreover, the large variance of the investment rate, which is an indicator of firm growth, suggests that unobserved heterogeneity is probably necessary to explain the observed investment patterns.

Although the features and patterns of the data just described point to some possible effects of the policy on firms' behavior, inference cannot be made from the raw data because firms are heterogeneous in unobservable productivity, therefore investment patterns may be a result of unobserved differences among subsidy recipients and non-recipients. To control for unobserved heterogeneity, in section 6 I estimate firm-specific productivity using input and output data and the economic model developed in section 4.

6 Estimation of Demand, Production Cost and Productivity

In this section I build on the implications of the economic model presented in section 4 to construct the econometric model and devise an estimation strategy to recover unobserved firm-specific productivity using firm-level panel data on revenue, labor cost and expenditures in materials. The estimated productivity is used both in the reduced form investment demand estimation in section 7 and in the estimation of the dynamic model in section 8. The econometric model developed in this section also recovers the parameters of the demand, marginal cost function, the productivity Markov process which will be used in the structural estimation of the dynamic model.

I begin by defining the econometric model for demand, marginal cost and productivity evolution. In what follows j enumerates firms and $t = 1, \ldots, T$ is an indicator of time. The log of the profit function in (3) is complemented with an *iid* error u_{jt} reflecting measurement error or optimization errors in choosing prices and becomes

$$\ln \pi_{jt} = \ln \left[-\frac{1}{\eta} \left(\frac{\eta}{\eta+1} \right)^{\eta+1} \Phi \right] + (\eta+1)(\beta_k \ln k_{jt} + \beta_w \ln w - \omega_{jt}) + u_{jt} \tag{10}$$

$$\pi_{jt} = -\frac{1}{\eta} \left(\frac{\eta}{\eta+1} \right)^{\eta+1} \Phi \left(e^{\beta_w \ln w} \right)^{\eta+1} \left(e^{\beta_k \ln k_{jt} - \omega_{jt}} \right)^{\eta+1} e^{u_{jt}} \tag{11}$$

It is important to stress that the purpose of the static estimation presented in this section is to recover the profit function parameters and the productivity evolution affecting investment choice and it is not to estimate the underlying marginal cost function and demand function themselves. Consequently, I have no interest in separately identifying Φ, β, β_w because, as equation (11) illustrates, quantities Φ and $\left(e^{\beta_0 + \beta_w \ln w}\right)^{\eta+1}$ do not vary across firms[16] and only affect the scale of the profit function. Thus, I gather all scale quantities together and rewrite (10) as

$$\ln \pi_{jt} = \ln(A) + (\eta+1)(\beta_k \ln k_{jt} - \omega_{jt}) + u_{jt}, \; u_{jt} \sim iid \tag{12}$$

In equation (12) the realization of variables π, k is observable but the realization of variables ω, u is not.

[16]The underlying assumption is that all firms face the same variable input prices w at any point in time.

Given that the marginal cost is assumed to be constant, using the definition of the total variable cost together with the optimal price rule in (2) and assuming *iid* optimization errors in the optimal price decision, the following relationship between total variable cost and revenue holds

$$\text{tvc}_{jt} = c_{jt}q_{jt} + \nu_{jt} = \frac{\eta+1}{\eta}p_{jt}q_{jt} + \nu_{jt} = \left(1 + \frac{1}{\eta}\right) r_{jt} + \nu_{jt},\ \nu_{jt} \sim iid \tag{13}$$

Assuming constant elasticity of demand and constant marginal cost implies that the optimal price equals marginal cost plus a markup directly related to the elasticity of demand η. Therefore, the relationship between total variable cost and revenue provides information on the parameter η which can be estimated directly from (13) with ordinary least squares as in Aw, Roberts and Xu (2011).

The productivity process is a non-linear first order Markov process with an *iid* normal innovation in productivity ξ specified as

$$\omega_{jt+1} = E\left(\omega_{jt+1}|\omega_{jt}\right) + \xi_{jt} = \rho_0 + \rho_1\omega_{jt} + \rho_2\omega_{jt}^2 + \xi_{jt},\ \xi_{jt} \overset{iid}{\sim} N(0, \sigma_\xi^2) \tag{14}$$

In order to consistently estimate F_ω and β_k it is necessary to be able to separate ω_{jt} from u_{jt}. Separating ω from u is also crucial because ω is the serially correlated shock profit function and dynamic decisions are based on ω but not on u. Based on the insights of Levinsohn and Petrin (2003) and Olley and Pakes (1996), to estimate β_k, F_ω it is sufficient to find an observable variable correlated with productivity ω such that , conditioning $(\eta+1)\omega_{jt} + u_{jt}$ on this variable, the existing statistical dependence between the observable $(\eta+1)\omega_{jt} + u_{jt}$ with the unobservable ω is eliminated. In datasets like the one used in this paper, possible candidates for such a variable are investment expenditures (as in Olley and Pakes (1996)), materials expenses or labor costs (as in Levinsohn and Petrin (2003)). Using investment would not be an appropriate choice in my case because investment demand depends on many factors, such as subsidy status, subsidy rate, expectations about future prices of capital, that differ across firms and are hard to control for without explicitly modeling firm's dynamic decision to invest. I choose the combined expenditures in materials and electricity as they rarely exhibit zero values in the data and are the most variable inputs. Because materials and electricity are variable inputs, it is reasonable to assume that they are chosen after ω is observed and hence are correlated with current productivity.

More specifically, I assume that productivity can be written a function of capital k and the total expenditure on materials and electricity, that is $\omega = \omega(k, m)$[17]. This implies that equation (12) can be written as

$$\begin{aligned} \ln \pi_{jt} =& \gamma + (\eta + 1)\beta_k \ln k_{jt} - (\eta + 1)\omega(k_{jt}, m_{jt}) + \tilde{u}_{jt} \\ \ln \pi_{jt} =& \gamma + h(k_{jt}, m_{jt}) + \tilde{u}_{jt}, \;\; E(\tilde{u}_{jt}|k, m) = 0 \end{aligned} \tag{15}$$

with function $h(\cdot)$ capturing the combined effect of capital and productivity on variable profits. I specify h as a second degree polynomial with interaction terms and estimate all the parameters in (15) with ordinary least squares. Given consistent estimates $\hat{\eta}, \hat{h}$, following Levinsohn and Petrin (2003) and Aw, Roberts and Xu (2011), I can construct the productivity series for every firm as a function of β_k by defining:

$$\tilde{\omega}(k, \beta_k) = (\hat{\eta} + 1)\beta_k \ln k - \frac{\hat{h}(k, m)}{\hat{\eta} + 1} \tag{16}$$

[17]For the derivation of $\omega(k, m)$ from a production function consistent with my demand and cost assumptions see appendix B.2.

and substituting (16) into (14) I have that

$$\begin{aligned} \tilde{\omega}(k_{jt+1}, \beta_k) = \rho_0 + \rho_1 \tilde{\omega}(k_{jt}, \beta_k) + \rho_2 \tilde{\omega}^2(k_{jt}, \beta_k) + \tilde{\xi}_{jt} \\ E(\tilde{\xi}_{jt} | k_{jt}, \omega_{jt}) = 0 \end{aligned} \tag{17}$$

I estimate $\beta_k, \rho_0, \rho_1, \rho_2$ from (17) with non-linear least squares and σ_ξ as the sample variance of $\tilde{\xi}$, $\hat{\sigma}$.Finally, given estimates $\hat{\eta}, \hat{\beta}_k, \hat{h}$ I can recover the productivity series $\{\hat{\omega}_{jt}\}_t$ for every firm j.

To recap, using data for every firm on total variable cost, total revenue, capital stock, materials and electricity expenditure I estimate demand elasticity η and the coefficient of the logarithm of capital in the marginal cost function β_k, which jointly characterize the profit function. Furthermore, I estimate the productivity series ω_{jt} for each firm which captures firm heterogeneity. Lastly, I recover the productivity transition function F_ω by estimating (g, σ_ξ). F_ω is used to calculate firms' expectations about future productivity which, in conjunction with the profit function and the current state, affect firms' investment choice.

6.1 Estimation results

The parameters of the marginal cost function, the demand function and the productivity process in equations (12) and (14) are reported in table 15. Demand elasticity η is -2.68 which results in a markup of price over marginal cost of 60 percent. The coefficient β_k associated with $\ln k$ in the marginal cost function is -0.29 indicating that larger firms have lower marginal cost. The elasticity of demand η and the marginal cost coefficient on capital β_k jointly determine the exponent of capital in the variable profit function π defined in equation (4) which has an estimate of .48. This value means that the π is an increasing concave function with respect to capital stock. The curvature of the variable profit function is one of the main forces driving firms' investment decisions because it affects the marginal benefit of investment and plays a crucial role in the estimation of the capital adjustment cost parameters. Parameters ρ_0, ρ_1, ρ_2 determine the conditional expectation in the productivity process defined in (14). Notice that the coefficient ρ_2 on the non-linear quadratic term is small (.0006) and statistically insignificant implying that an AR(1) specification could be a good approximation of the productivity evolution. The coefficient ρ_1 (.99) indicates a very high serial correlation in ω_t. This highly persistent productivity process suggests that productivity is a very important factor behind firms' investment choices because today's productivity is a very good predictor of future productivity and, consequently, future profitability. The mean of the productivity process is .35 and the 10th and 90th percentiles are -.13 and .96, respectively. These figures signify that a firm in the 90th percentile of the productivity distribution has variable profits roughly 6 times higher than a firm of similar size in the 10th percentile. The high persistence and cross-sectional variation of productivity demonstrate that controlling for it is key in order to make correct inferences.

7 Reduced-form estimation of the subsidy policy treatment effect

In this section I use the observed investment decisions to analyze the effect of the subsidy policy on firm investment. I first show how the behavior of firms changes when they are granted a subsidy and then show the effect of the policy on unsubsidized firms exploiting the variation of the intensity of the policy across locations and time.

The dynamic model of investment developed in section 4 implies that the investment demand equation is a function of the form $i = i(\omega, p_k, k, l, \epsilon; F_\omega, F_{p_k}, F_\epsilon)$. The estimation method of the labor demand equation depends on the research question of interest. For example, if we are interested in quantifying the difference in investment demand between subsidized and unsubsidized firms under the current policy regime, it is not necessary to estimate the effect of the policy regime itself F_{p_k} on firm investment demand. A projection of $i(\omega, p_k, k, l, \epsilon; F_\omega, F_{p_k}, F_\epsilon)$ on the space spanned by (ω, p_k, k, l) suffices. Such a projection would be policy variant in the sense that the estimated relationship between (ω, p_k, k, l) and i would change if the policy regime F_{p_k} where to change. If we want to simulate investment demand under different policy regimes F_{p_k} as in section 8, a policy-invariant (structural) relationship between (ω, p_k, k, l) and i should be estimated. The econometric models aiming to answer each research question arising in this paper share the following basic structure

$$\pi = \pi(k, \omega) \tag{Profit Function}$$

$$\omega_{t+1} = E\left(\omega_{t+1}|\omega_t\right) e^{\xi_t} \tag{Productivity}$$

$$\left.\begin{array}{l} V(\omega, p_k, k, l, \epsilon) = \max\limits_{i \geq 0} \pi(\omega, k) - (p_k i + \dfrac{c_2}{2}\dfrac{i^2}{k})\epsilon + \\ +\beta \displaystyle\int_{\omega', p_k', \epsilon'} V(\omega', p_k', k', l, \epsilon) dF_\omega(\omega'|\omega) dF_{p_k}(p_k'|p_k, l, \omega) dF_\epsilon(\epsilon') \\ \Leftrightarrow \\ i = i(\omega, p_k, k, l, \epsilon; \pi, c_2, F_\omega, F_{p_k}, F_\epsilon) \end{array}\right\} \tag{Investm.}$$

The one period time-to-build assumption along with the exogenous evolution of productivity render the sub-system of equations {(Profit Function),(Productivity)} autonomous which implies that the profit function and the productivity process can be estimated in a first stage which is done in section 6.

The goal of this section is to estimate the average treatment effect of the policy that is

$$\text{ATE} = E\left\{i(\omega, p_k, k, l, \epsilon; \cdot | p_k < 1) - i(\omega, p_k, k, l, \epsilon; \cdot | p_k = 1)\right\} \tag{18}$$

Specifically, in this section I estimate the average treatment effect (ATE) of the policy using a reduced form of the optimal investment function where the investment equation does not explicitly depend on the objects $\epsilon, F_\omega, F_{p_k}, F_\epsilon$ in the way implied by the Bellman equation (9). Instead, I use location and year fixed effects to approximate the differential effect of the policy in different locations and time periods. I separate Greek regions in two broad categories depending on the intensity of the subsidy policy. In locations belonging to category H the subsidy program is particularly generous both in terms of subsidy rates and allocated funds. Consequently, it is easier for firms in H to be subsidized and, if they are subsidized, to benefit from high subsidy rates. Regions not in H belong to location category L. The categorization of regions in H and L is presented in table 6. Table 7 shows the different intensity of the policy in the two locations. In location H 44 percent of the firms applied for a subsidy and 39 percent of those were eventually subsidized at some point during the sample period bringing the odds of being subsidized, conditional on applying for the subsidy, to 88 percent. In contrast, in location L 29 percent of the firm applied for a subsidy and 24 percent of those were eventually subsidized at some point during the sample period bringing the odds of being subsidized conditional on applying to 83 percent. Note that the high odds of being subsidized conditional on applying imply that firms self-select into applying either because of high application costs or for other reasons that are beyond the scope of this paper. The average waiting time in days between the subsidy application and the successful[18] result is 253 for H and 195 for L which means

[18] In the data I do not observe the waiting time for a rejection because I can only observe the application date and the subsidy grant decision date. I consider the waiting period of successful applicants as a proxy for the waiting period for unsuccessful applicants as well.

that firms might have high incentives to postpone their investment plans for a about a year anticipating the outcome of the application process. The median firm in location H is similar in terms of capital stock and number of employees to the median firm in L but firms located in H are more capital intensive. This can be due to the policy which effectively decreases the price of capital relative to the price of labor. I use the categorization of the location to define the dummy variable D^l_{jt} which takes the value 1 if firm j is located in H at year t. This dummy variable is used to proxy the indirect effect of location on firms' investment decisions.

As mentioned before, I can observe the year a firm is granted a subsidy but not for how long the firm is subsidized; that is, the duration of the subsidy spell is unobservable. Table 4 shows that the median investment rate and investment level drop 4 years after a firm has been granted a subsidy indicating that firms are subsidized only for a limited number of years even if extensions are possible. I define dummy variables $D^2_{jt}, D^3_{jt}, D^4_{jt}$ which are equal to 1 if a firm was granted a subsidy within the last 2, 3 or 4 years, respectively. Each dummy variable identifies different sets of observations as the treated population and the results from alternative specifications of the model (where each specification depends on how the treated population is identified by the dummy variables) can be compared to test the robustness of the ATE to the definition of the treated population.

Specifically, the estimating equation is an investment equation of the intensive margin with 3 different specifications, one for each of the treatment dummy D^{sp} where $sp = 2, 3, 4$.

$$\begin{aligned}\ln i =&\gamma_{ATE} D^{sp}_{jt} + \gamma_L D^l_{jt} + \sum_{t=1}^{T} \gamma_t d^t + \gamma_k \ln k_{jt} + \gamma_{kk} (\ln k_{jt})^2 + \\ &+ \gamma_\omega \omega_{jt} + \gamma_{\omega\omega} (\omega_{jt})^2 + \gamma_{k\omega} \omega_{jt} \ln k_{jt}, \;\; i > 0 \end{aligned} \tag{19}$$

Table 8 reports the results from the estimation of equation (19). Note that the coefficient associated with the treatment dummy is statistically and economically significant in all specifications. The policy induces subsidized firms to invest, on average, approximately 60 percent more than unsubsidized firms *ceteris paribus*.

To verify that the effect of the treatment is positive, significant and robust to different specifications I also estimate an investment demand equation of the form

$$i = \gamma_{ATE} D_{jt}^{sp} + \gamma_L D_{jt}^{l} + \sum_{t=1}^{T} \gamma_t d^t + \gamma_k k_{jt} + \gamma_\omega \exp(\omega_{jt}) + \gamma_{k\omega} \exp(\omega_{jt}) k_{jt} \tag{20}$$

The investment demand equation in (20) is estimated with a Tobit likelihood where both the right hand side and the left hand side are in levels. The results from this alternative estimation are reported in table 9. Once again the coefficient of the treatment dummy is statistically and economically significant in all specifications. The policy induces subsidized firms to invest, on average, approximately .3 million Euros more than unsubsidized firms, *ceteris paribus*.

Another quantity of interest is the effect on the investment rate which is a proxy for the growth rate of the firm. Thus, I estimate the following equation to evaluate the average treatment effect of the policy on the investment rate

$$i/k = \gamma_{ATE} D_{jt}^{sp} + \gamma_L D_{jt}^{l} + \sum_{t=1}^{T} \gamma_t d^t + \gamma_k \ln k_{jt} + \gamma_\omega \omega_{jt} + \gamma_{k\omega} \omega_{jt} \ln k_{jt} \tag{21}$$

The estimation of equation (21) is performed via Tobit likelihood and the results are reported in table 10. Also in this case the coefficient of the treatment dummy is statistically and economically significant in all specifications. The policy induces subsidized firms to exhibit, on average, growth rates (in terms of investment rate) approximately 1 percent higher than unsubsidized firms, *ceteris paribus*.

The way the policy affects aggregate investment is of particular interest here. To understand how the policy might influence aggregate investment I finally try to quantify the effect of the subsidy policy on the probability of an investment spike. An investment spike is defined as an investment rate above 20 percent (see Cooper and Haltiwanger (2006)) and is of particular relevance in this context as most aggregate investment occurs in investment spikes. I define a spike dummy variable D^{spike} taking the value 1 for observations with investment rate greater than 20 percent. Table 11 shows that even though only a quarter of the observations are spikes, these investment instances are responsible for approximately 60 percent of aggregate investment. The effect of the policy on the probability of a spike is evaluated using the following probit model

$$\begin{aligned}\text{Prob}(D^{\text{spike}} = 1|\cdot) = \Phi \Bigg(& \gamma_{ATE} D^{sp}_{jt} + \gamma_L D^{l}_{jt} + \sum_{t=1}^{T} \gamma_t d^t + \gamma_k \ln k_{jt} + \gamma_{kk} (\ln k_{jt})^2 + \\ & + \gamma_\omega \omega_{jt} + \gamma_{\omega\omega} (\omega_{jt})^2 + \gamma_{k\omega} \omega_{jt} \ln k_{jt} \Big) \end{aligned} \tag{22}$$

where Φ is the standard normal cdf. Table 12 reports the results of the maximum-likelihood probit estimation. Also for investment spikes the main result is that the coefficient of the treatment dummy is statistically and economically significant in all specifications with the subsidy policy inducing subsidized firms to undertake investment spikes more often than unsubsidized firms, *ceteris paribus.*

7.1 Evidence of the effect of the policy on unsubsidized firms

The results from the estimation of the treatment effect demonstrate that the investment behavior of firms that receive a subsidy at some point during the sample period is very different than the investment behavior of firms that are never subsidized and this different behavior is observable also in the years in which subsidized firm are not directly benefiting from the subsidy. To this extent, it would be interesting to verify how the behavior of firms in the presence of the policy changes in comparison to an environment where there is no subsidy policy. This is similar to what Heckman and Vytlacil (2007) call a policy relevant treatment effect (PRTE). It would be also interesting to decompose the PRTE into the effect of the policy on the investment behavior during the subsidized years (policy relevant treatment effect on the treated - PRTET) and the effect of the policy on the investment behavior of firms that are never subsidized and always face a market price of capital of $p_k = 1$ (policy relevant treatment effect on the untreated - PRTEU). Define PRTE, PRTET and PRTEU as

$$\begin{aligned}
\text{PRTE} =& E\left\{i(\omega, p_k, k, l, \epsilon; F_\omega, F_{p_k}, F_\epsilon) - i(\omega, p_k = 1, k, l, \epsilon; F_\omega, F'_{p_k}, F_\epsilon)\right\} \\
\text{PRTET} =& E\left\{i(\omega, p_k < 1, k, l, \epsilon; F_\omega, F_{p_k}, F_\epsilon) - i(\omega, p_k = 1, k, l, \epsilon; F_\omega, F'_{p_k}, F_\epsilon)\right\} \\
\text{PRTEU} =& E\left\{i(\omega, p_k = 1, k, l, \epsilon; F_\omega, F_{p_k}, F_\epsilon) - i(\omega, p_k = 1, k, l, \epsilon; F_\omega, F'_{p_k}, F_\epsilon)\right\}
\end{aligned}$$

where $F'_{p_k}(p_k = 1) = 1$ identifies an environment without the subsidy policy. If the outcome of interest is the result of a dynamic decision that is affected by the policy, as the result of the decision to invest in this paper, PRTEU is not zero because the presence of the policy affects firms' expectations, and consequently their decisions, even in states where they are not treated, i.e. in states where they are not receiving the subsidy. This implies that in general, PRTE$\neq ATE$.

The purpose of this section is to test whether the effect of the policy on unsubsidized firms is quantitatively important. Unfortunately, a state where $F'_{p_k}(p_k = 1) = 1$ is never observed in the data because the subsidy policy is always in place during the sample period so the PRTEU is not observable. On the other hand, the intensity of the policy dramatically changed after 2004, thus this structural change can be exploited to evaluate whether the behavior of unsubsidized firms changed accordingly after 2004. Table 13 presents statistics about the intensity of the policy before and after 2004. The share of total cash subsidies allocated after 2004 is 92 percent of the total cash subsidies allocated during the entire sample period and the share of the number of grants allocated after 2004 is 85 percent. The median subsidy rate jumps from .3 to .4 and the waiting period for a successful application process almost doubles. These statistics show that the scope of the policy widened so relevantly after 2004 that it very likely to assume that unsubsidized firms started to take into account the possibility of receiving a subsidy in the future when making investment decisions. Moreover, the effect is expected to be different in location H and location L since the potential gains from a subsidy are higher for a firm in location H than for on in location L. I exploit this differential effect of the policy in different locations applying a difference in difference design to evaluate the quantitative importance of the policy on unsubsidized firms.

Let D^{after} be a dummy variable taking the value of 1 for the years after 2004. The baseline estimating equation is given by

$$\begin{aligned}\ln i =&\gamma_{DID} D_{jt}^{after} D_{jt}^{l} + \gamma_{after} D_{jt}^{after} + \gamma_L D_{jt}^{l} + \gamma_k \ln k_{jt} + \gamma_{kk} (\ln k_{jt})^2 + \\ &+ \gamma_{\omega}\omega_{jt} + \gamma_{\omega\omega}(\omega_{jt})^2 + \gamma_{k\omega}\omega_{jt} \ln k_{jt},\ i > 0,\end{aligned} \tag{23}$$

The coefficient of interest is γ_{DID} which accounts for the differential effect on firms located in H after the policy change. In this context it is important to understand what is the appropriate subsample to estimate (23) in order to capture the dynamic effect of the policy on unsubsidized firms. I choose to exclude from the sample all the observations from the year a firm receives a subsidy onwards. This choice is dictated by the fact that I do not want γ_{DID} to capture the incentive for subsidized firms to over-invest when they are subsidized and under-invest in the future; that is, the inter-temporal substitution effect. Let t_{j0} be the year firm j receives a subsidy in the sample for the first time and set $t_{j0} = \infty$ for firms that never receive a subsidy in the sample. I exclude from the estimation sample the observations i_{jt} for which $t \geq t_{j0}$. The theoretical model also indicates that the unsubsidized firms whose investment behavior is affected by the policy are those believing that with positive probability they will be subsidized in the future. Even if firm's idiosyncratic expectations are unobservable and cannot be conditioned on, the data on applications allow for separating firms into those that apply for a subsidy at some point during the sample period and those that never apply. Let t_{ja} be the year a firm applies for a subsidy and set $t_{ja} = \infty$ if firm j never applies for a subsidy during the sample period. Then, the sample of applicants can be defined as the observations i_{jt} for which $\exists\, t < \infty \ni t = t_{ja}$. I estimate (23) for these three different samples

$$
\begin{aligned}
\text{Sample Unsub. } \mathbf{U} = &\quad i_{jt} \ni t < t_{j0} \\
\text{Sample Unsub. Applicants } \mathbf{U}_a = &\quad i_{jt} \ni t < t_{j0},\ \exists\, t < \infty \ni t = tja \\
\text{Sample Unsub. Non-Applicants } \mathbf{U}_{na} = &\quad i_{jt} \ni t \geq t_{j0}, t_{ja} = \infty \\
\Rightarrow \mathbf{U}_{na} \cup \mathbf{U}_a = \mathbf{U}, &\quad \mathbf{U}_{na} \cap \mathbf{U}_a = \varnothing
\end{aligned}
$$

The results from the estimation of (23) for the three samples $\mathbf{U}_{na}, \mathbf{U}_a, \mathbf{U}$ are reported in table 14. In the unsubsidized applicants sample, the diff-in-diff coefficient is statistcally significant and negative implying that unsubsidized firms that will eventually apply for a subsidy on average decrease their investment approximately by 50 percent due to the increase in the intensity of the policy, *ceteris paribus*. This result is an indication that there is a quantitatively important effect of the policy on unsubsidized firms, probably prompted by the dynamic nature of the capital input. In the unsubsidized non-applicants sample the diff-in-diff coefficient becomes insignificant. It is possible that, even though the policy expanded in scope after 2004, the firms in this sample do not expect to ever receive a subsidy and therefore do not change their behavior because of the policy. In the combined sample of unsubsidized applicants and non-applicants the diff-in-diff coefficient is also insignificant. Probably this result depend on the fact that the sample size of non-applicants is twice as large as the sample of applicants. The last two rows of table 14 show that the unsubsidized applicants' share of total investment and sales is 47 and 39 percent, respectively. This implies that if the quantities of interest are aggregate investment or aggregate productivity, the effect of the policy on unsubsidized applicants is large enough to affect aggregate quantities as their investment and sales share are substantial despite their limited presence in the sample.

7.2 Conclusions from the reduced-form analysis

In this section I specify and estimate a reduced form investment demand equation to quantify the effect of the subsidy on investment demand. My results show that subsidized firms changed their investment demand substantially because of the subsidy policy. Moreover, using a change in the intensity of the policy in the middle of the sample and the asymmetric implementation of the policy across locations I evaluate the effect of the policy on the investment demand of unsubsidized firms. I find that unsubsidized firms which will eventually apply for the policy at some point during the sample period, substantially reduce their investment demand while unsubsidized firms which will never apply for a subsidy do not exhibit any change in investment demand. This heterogeneous effect of the policy on applicant and non-applicant unsubsidized firms suggests that the expectation of firms about the possibility of receiving a subsidy in the future affects their investment behavior in the current period and this phenomenon can have an impact on aggregate outcomes. In order to evaluate the effect of the policy on aggregate outcomes and decompose the heterogeneous impact of the policy on unsubsidized and subsidized firms, I develop and estimate a fully specified structural econometric model in section 8.

8 Structural Estimation

This section extends the dynamic model of investment developed in section 4 to a structural econometric model of investment in order estimate the capital adjustment cost parameters characterizing $C(\cdot)$ and the distribution F_ϵ of the idiosyncratic investment cost ϵ. The econometric model is structural in the sense that its parameters are policy invariant; that is, their values do no change if the subsidy policy changes in intensity or ceases to exist altogether. This policy-invariant feature of the parameters permits the simulation of the firms' choices in a counterfactual environment never observed in the data such as an economy without the subsidy policy. Since the policy is active during all the years included in the dataset, to estimate the parameters of the model it is necessary to account for the effect of the policy on the observed investment behavior by specifying and estimating the subsidy policy rule characterized by the transition of the price of capital F_{p_k}. Certain policy invariant parameters of the model, that is the profit function and the Markov productivity process, are estimated in section 6 and are considered known in the econometric model presented this section. The firm-specific productivity, also estimated in section 6, is considered given as well. The counterfactual simulations performed after the dynamic estimation quantify the effect of the policy on all subsidized firms and the heterogeneous effect of the policy on low productivity and high productivity unsubsidized firms. The ability to perform counterfactual simulations of firm behavior in the absence of the policy is the main advantage of the structural dynamic model which is not an option in the reduced-form setup of section 7.

8.1 Specification of the subsidy policy rule

In this section I specify a functional form for $F_{p_k}(p'_k|p_k, l, \omega)$, the transition function of the price of capital, which summarizes *all* the information about the policy needed for a firm to form expectations about the future price of capital. Unsubsidized firms face price of capital $p_k = 1$ while subsidized firms face price $p_k < 1$ that depends on the subsidy rate. For expositional purposes, it is convenient to break $F_{p_k}(p'_k|p_k, l, \omega)$ in two components: component $F^S_{p_k}$ which characterizes the expectations of currently subsidized firms about next period's price of capital, and component $F^U_{p_k}$ which characterizes the expectations of currently unsubsidized firms about next period's price of capital.

$$\begin{aligned} F^S_{p_k} &:= F_{p_k}(p'_k|p_k, l, \omega),\ p_k < 1 \\ F^U_{p_k} &:= F_{p_k}(p'_k|p_k, l, \omega),\ p_k = 1 \end{aligned} \tag{24}$$

8.1.1 The distribution of future price of capital for subsidized firms $F^S_{p_k}$

I assume that each currently subsidized firm will face the same subsidy rate next period with probability λ_k and zero subsidy rate with probability $1 - \lambda_k$ independently of firm's productivity or location. Formally, subsidized firms are assumed to face a two point support discrete probability distribution of the form

$$F^S_{p_k}(p'_k|p_k) = \begin{cases} \lambda_k & \text{if } p'_k = p_k \\ 1-\lambda_k & \text{if } p'_k = 1 \\ 0 & \text{otherwise} \end{cases} \tag{25}$$

This assumption is made partly to match some institutional characteristics of the policy and partly to avoid overcomplicating the model. Even though being subsidized for two consecutive years with a different subsidy rate is possible, it is highly unlikely and I never observe it in the data, therefore the two point support assumption underlying (25) appears to be justified.

Typically, a subsidized firm has a pre-specified number of years to undertake a specific investment expenditure with a specified subsidy rate. The number of years are known to the firm but unknown to the econometrician. Extensions to both the deadline to complete the investment project and the expenditure size are possible but uncertain at the time the subsidy is assigned. Reducing the characterization of the subsidy policy to a subsidy rate $1 - p_k$ is a simplification driven by the necessity to keep the number of state variables to a minimum[19].

8.1.2 The distribution of future price of capital for unsubsidized firms $F^U_{p_k}$

Unsubsidized firms face probability distribution $F^U_{p_k}(p'_k|p_k, l, \omega)$ which depends on both their location and their current productivity. $F^U_{p_k}$ depends on location because the subsidy policy analyzed here is location-dependent and also depends on firm productivity to account for both the self-selection of firms into the policy and a possibly productivity-dependent government allocation rule. By design of the subsidy policy, the subsidy rate takes values in a discrete set[20] so the support of the price of capital is the finite set $\Delta = \{p_{k1}, \ldots, 1\}$. There is a finite number of locations which means that the location variable belongs to a finite set $\Lambda = \{l_1, \ldots, l_{N_\Lambda}\}$. I decompose $F^U_{p_k}(p'_k|\omega, l)$ in two components: the probability of receiving a subsidy P_S and a location-dependent probability mass function of subsidy rates conditional of receiving a subsidy $G(\cdot, l)$ that is

[19]Incorporating both the project size and the firm's decision to apply for the subsidy program into the model is ongoing research.

[20]All subsidy rates are multiples of .05

$$F^U_{p_k}(p'_k|1,l,\omega) = \begin{cases} P_S(l,\omega)G(p'_k,l) & ,\text{if } p'_k < 1 \\ 1 - P_S(l,\omega) & ,\text{if } p'_k = 1 \end{cases} \tag{26}$$

$$P_S(l,\omega) = exp(\beta^l_{p0} + \beta^l_\omega \omega)/(1 + exp(\beta^l_{p0} + \beta^l_\omega \omega)),\ l \in \Lambda \tag{27}$$

Where $P_S(l,\omega)$ is the probability a currently unsubsidized firm in location l with productivity ω will receive a subsidy next period and $G(p'_k,l)$ is the probability a subsidized firm in location l will face price of capital p'_k. This distribution implies that the expected next period's price of capital for a currently unsubsidized firm in location l with productivity ω is $1 - P_S(l,\omega) + P_S(l,\omega)\sum_{p'_k<1}^{G}(p'_k,l)p'_k$. Productivity affects only the probability of firms being subsidized while the subsidy rate/price of capital for subsidized firms is independent of productivity and depends only on the location of the firm. I use a logit specification for the probability of receiving a subsidy conditional on productivity ω.

To summarize, I model the subsidy policy rule as a Markov chain of the price of capital which is equal to one for unsubsidized firms and less than one (equal to one minus the subsidy rate) for subsidized firms. In each period, unsubsidized firms face a probability distribution over future prices of capital which depends on their location and productivity. This dependence of productivity controls for the self-selection of the most productive firms into the policy which induces heterogeneity in expectations about the future price of capital. Subsidized firms expect tomorrow's price of capital to be equal to today's price with probability λ_k or equal to one with probability $1 - \lambda_k$.

8.2 Specification of adjustment cost $C(\cdot)$ and the distribution F_ϵ of the structural error

I specify $C(k,k') = \frac{c_2}{2}\frac{(k'-(1-\delta)k)^2}{k}$ as a quadratic adjustment cost function jointly convex in investment $(k'-(1-d)k)$ and capital if $c_2 \geq 0$. High values of c_2 in this specification prevent firms to exhibit hight growth rates even if they experience very high productivity shocks. This form of convex adjustment cost is common in the literature[21] and it implies that the marginal cost of adjustment is linear in the ratio of investment to capital. Several papers[22] add a fixed cost component in the specification of the adjustment cost function. One of the reasons they do so is to rationalize the presence of a probability mass at zero investment since a considerable number of observations have zero investment in most datasets. In my model, the complete irreversibility of investment is enough to rationalize a probability mass at zero investment. An added benefit of a convex adjustment cost specification is that, in combination with a concave profit function in capital, it delivers a concave value function in capital which can be exploited in the numerical solution of the maximum likelihood problem.

The distribution of the structural error ϵ is assumed to be lognormal i.e. $\ln(\epsilon) \sim N(0,\sigma_\epsilon^2)$. I refer to ϵ as structural because ϵ has a clear interpretation in the context of the model and firms form expectations about the future values of ϵ when making investment decisions. The choice of a mean-zero normal distribution for $\ln(\epsilon)$ is not an innocuous assumption in the model and affects the interpretation of the adjustment cost parameter c_2. Under my specification, $\frac{c_2}{2}\frac{(k'-(1-\delta)k)^2}{k}$ is the adjustment cost faced by a firm with ϵ equal

[21] The same specification of convex adjustment costs is assumed in Asker, Collard-Wexler and De Loecker (2014); Cooper and Haltiwanger (2006); Fuentes, Gilchrist and Rysman (2006); Gilchrist and Himmelberg (1995).

[22] For example Asker, Collard-Wexler and De Loecker (2014); Cooper and Haltiwanger (2006); Fuentes, Gilchrist and Rysman (2006)

to the median of the distribution F_ϵ[23]. The variance σ_ϵ^2 is an integral component of the solution of the firm's dynamic programming problem and is crucial in making probabilistic statements about the outcome of counterfactual policy scenarios.

8.3 Measurement issue: the partial observability of variable p_k

This section clarifies the problem of partial observability of the price of capital p_k for the econometrician and describes how this problem is dealt with in the estimation. The partial observability arises because in the data only the year a firm is granted a subsidy is observed but not for how long the firm is subsidized; that is, the duration of the subsidy spell is unobservable. Consequently, the partial observability of the variable p_k does not arise because the subsidy rate is observed with error but because it is unknown (by the econometrician) whether a firm which received a subsidy grant at time t_0 with subsidy rate $1 - p_{k0}$ is still subsidized at a later time $t_0 + r,\ r > 0$. This phenomenon can be thought as a peculiar form of measurement error . Peculiar in the sense that in a subsample of the data, that is for firm-year observations corresponding to years after a firm is granted a subsidy, p_k is observed with measurement error while for a subsample of the data, that is for firm-year observations corresponding either to years before a firm was granted a subsidy or to the observations of firms never subsidized in the sample, p_k is observed without measurement error. Moreover, in the subsample in which p_k is not observed, the econometrician knows that p_k can only take two possible values: either value 1, meaning that firm is not subsidized anymore, or the value p_{k0} corresponding to the investment rate $1 - p_{k0}$ with which the firm was granted the subsidy. Consequently, under the assumption that the subsidy spell is independent to all other state variables, the unobserved p_k can be integrated out of the criterion function in the estimation as long as the econometrician

[23]If the specification was $\ln(\epsilon) \sim N(-\frac{\sigma_\epsilon^2}{2}, \sigma_\epsilon^2)$ then $\frac{c_2}{2}\frac{(k' - (1-\delta)k)^2}{k}$ would have been the adjustment cost faced by a firm with ϵ equal to the mean of the distribution F_ϵ.

can put a probability on the event $(p_k = p_{k0},\ t = t_0 + r)$. The specification of F_{p_k} implies that the probability a firm which received a subsidy grant at time t_0 is still subsidized at $t_0 + r$ is λ_k^r.

This immediately raises the issue of how to estimate λ_k since the econometrician never observes the duration of the subsidy spell. The solution is to estimate λ_k jointly with the rest of the dynamic parameters. The identification of λ_k comes from the fact that p_k is unobservable only for a subsample of observation. Simply put, the rest of the dynamic parameters can be identified in the subsample where p_k is observable and observed firms' investment behavior can identify λ_k. Table 4 shows subsidized firms' investment patterns each year following the subsidy grant. Observe that the median investment level and investment rate increase right after the grant of the subsidy and then drops at the pre-subsidy level and even lower than the pre-subsidy investment rate. The sharp decrease in investment few years after the subsidy grant identifies λ_k.

8.4 Estimator and estimates of $F^U_{p_k}$

Recall from section 24 that the subsidy policy rule has two components: component $F^U_{p_k}$ which characterizes the expectations of currently unsubsidized firms about next period's price of capital, and component $F^S_{p_k}$ which characterizes the expectations of currently subsidized firms about next period's price of capital and is itself fully characterized by parameter λ_k. Section 8.3 lays out the reasons why λ_k, and consequently $F^S_{p_k}$, is estimated jointly with the rest dynamic parameters. This section describes the estimation of $F^U_{p_k}$ prior to the estimation of the dynamic parameters. By construction (see (26)) $F^U_{p_k}$ is decomposed in two components $P_S(l, \omega)$ and $G(p'_k, l)$. In this paper I estimate a location independent policy rule F_{p_k} in order to reduce the dimensionality of the dynamic programming problem but an extension to incorporate location in the empirical dynamic model is straightforwardThe probability of an unsubsidized firm to receive a subsidy $P_S(\omega) = exp(\beta_{p_0} + \beta_\omega \omega)/(1 + exp(\beta_{p_0} + \beta_\omega \omega))$ is specified as a logit probability and is estimated with maximum likelihood using the estimated firm-level productivity ω_{jt} from the estimation of the profit function in section 6 and the data on subsidy grants. The probability mass function of the subsidy rates $G(p_k')$ is estimated directly from the observed subsidy rates using frequencies.

The logit estimates are presented in tables 18 and the estimates of G are presented in 19. The coefficient on productivity in the logit model is positive indicating that more productive firms have higher probability of receiving a subsidy than unproductive firms. Figure 4 plots the predicted probabilities to receive a subsidy from the logit model versus productivity along with the productivity density. The model predicts that a firm in the highest decile of the productivity distribution has a 7 percent probability of receiving a subsidy which is 3.5 times higher than a firm in the lowest decile. Table 19 depicts the distribution of prices of capital for subsidized firms. Notice that the distribution of subsidy rates is skewed towards high rates with the highest subsidy rate (50 percent) being roughly twice as probable to be granted as any other subsidy rate. The estimated $F^U_{p_k}$ is considered as given in the estimation of the dynamic parameters.

8.5 Estimation of the dynamic parameters

The dynamic parameters of the model, i.e. c_2 which defines the capital adjustment cost function, the persistence in the subsidy policy rule λ_k and the variance σ^2_ϵ of the distribution of the idiosyncratic *iid* shock ϵ are estimated using a maximum likelihood estimator. Cost parameters in dynamic models of investment are typically estimated in two different ways. One possibility consists of using some version of firms' optimality condition such as the Euler equation in the vein of Hansen and Singleton (1982), or the first order condition derived from Q theory of investment as in Hayashi (1982). Alternatively, the indirect inference methods developed by Gourieroux, Monfort and Renault (1993) can be applied as in Cooper and Haltiwanger (2006) and Asker, Collard-Wexler and De Loecker (2014), where parameters are chosen such that simulated moments from the model are as close as possible to moments calculated from the data[24]. I take the first approach and use the first order necessary conditions of firms' dynamic programming problem.

For the first order conditions to be well defined, the expected value function of the firm should be differentiable with respect to capital. Recall that the dynamic problem of the firm is defined in (9). For a class of models like (9) where investment is completely irreversible Rincón-Zapatero and Santos (2009) prove that V is differentiable everywhere, even at the boundary of the feasible set, which implies that under regularity conditions, EV is also differentiable.

By dropping location l as a state variable the problem in (9) becomes

$$\begin{aligned} V(\omega, p_k, k, \epsilon) = \quad & \max_{i \geq 0} \pi(\omega, k) - \left[i p_k + \frac{c_2}{2} \frac{i^2}{k} \right] \epsilon + \\ & + \beta \int_{\omega', p_k', \epsilon'} V(\omega', p_k', k', \epsilon') dF(\omega', p_k' | \omega, p_k) dF_\epsilon(\epsilon') \\ & F(\omega', p_k' | s) = F_\omega(\omega' | \omega) F_{p_k}(p_k' | p_k, \omega) \\ & k' = i + (1-\delta)k, \ c_2 > 0 \end{aligned} \tag{28}$$

$$\Leftrightarrow \begin{aligned} V(s) = \quad & \max_{i \geq 0} \pi(\omega, k) - \left[i p_k + \frac{c_2}{2} \frac{i^2}{k} \right] \epsilon + \beta E \left[V(s') | s_{-\epsilon}, i \right] \\ & k' = i + (1-\delta)k, \ s := (\omega, p_k, k, \epsilon), \ s_{-\epsilon} := (\omega, p_k, k) \end{aligned} \tag{29}$$

Note that the expectation of next period's value function in (29) is conditional on $s_{-\epsilon}$ and not on s because ϵ is *iid*. This *iid* assumption simplifies considerably the calculation of the likelihood. Under differentiability of EV the first order necessary conditions for the

[24] For a comprehensive discussion of the methods suitable for the analysis of dynamic models of investment see Adda and Cooper (2003), Bond and Van Reenen (2007), and Pakes (1994).

optimal investment choice at state s are

$$\left.\begin{array}{l} -\left[p_k + c_2 \frac{i}{k}\right]\epsilon + \beta E\left[\frac{\partial V(s')}{\partial k'}\middle| s_{-\epsilon}, i\right] = 0 \quad \text{if } i > 0 \\ -\left[p_k + c_2 \frac{i}{k}\right]\epsilon + \beta E\left[\frac{\partial V(s')}{\partial k'}\middle| s_{-\epsilon}, i\right] < 0 \quad \text{if } i = 0 \end{array}\right\} \tag{30}$$

Rearranging terms and taking logs (30) becomes

$$\left.\begin{array}{l} \ln(\epsilon) = \ln\left(\beta E\left[\frac{\partial V(s')}{\partial k'}\middle| s_{-\epsilon}, i\right]\right) - \ln\left(p_k + c_2 \frac{i}{k}\right) \quad \text{if } i > 0 \\ \ln(\epsilon) > \ln\left(\beta E\left[\frac{\partial V(s')}{\partial k'}\middle| s_{-\epsilon}, i\right]\right) - \ln\left(p_k + c_2 \frac{i}{k}\right) \quad \text{if } i = 0 \end{array}\right\} \tag{31}$$

If it is assumed that p_k is always observed, the variables in the right hand side of (31) are observable; that is, for every observation in the data $x_{jt} := (k_{jt}, p_{k_{jt}}, \omega_{jt}, i_{jt})$, given EV, c_2, the log of the structural error $\ln(\epsilon)$ can be calculated directly. Given the parametric form of the distribution F_ϵ and letting f_ϵ be the density of $\ln(\epsilon)$ and $\theta := (c_2, \lambda_k, \sigma_\epsilon^2)$, the contribution of observation x_{jt} to the likelihood is given by

$$\begin{array}{c} L(x_{jt}|\theta) = \left\{ f_\epsilon\left(\ln\left[\frac{\beta E\left(\frac{\partial V(s')}{\partial k'}\middle| x_{jt}\right)}{p_k + c_2 \frac{i_{jt}}{k_{jt}}}\right]\right)\right\}^{d_{jt}} \left\{1 - F_\epsilon\left(\ln\left[\frac{\beta E\left(\frac{\partial V(s')}{\partial k'}\middle| x_{jt}\right)}{p_k + c_2 \frac{i_{jt}}{k_{jt}}}\right]\right)\right\}^{1-d_{jt}} \\ f_\epsilon, F_\epsilon = \text{pdf, cdf of } N(0, \sigma_\epsilon^2), \quad d_{jt} = \left.\begin{array}{ll} 1 & \text{, if } i_{jt} > 0 \\ 0 & \text{, otherwise} \end{array}\right\} \end{array} \tag{32}$$

which is the likelihood of a Tobit model since, in the case of zero investment, any $\epsilon > \beta E\left(\frac{\partial V(s')}{\partial k'}\middle| x_{jt}\right)/p_k + c_2 \frac{i_{jt}}{k_{jt}}$ can rationalize observation x_{jt}.

8.5.1 Integrating unobserved p_k out of the likelihood

For the reasons explained in section 8.3, for certain firm-year observations in the data p_k is a partially observed state variable and needs to be integrated out of the likelihood function. Recall from section 8.3 that to integrate out the partially observed $p_{k_{jt}}$ the econometrician needs to know the most recent year a firm received a subsidy which is defined as τ_{jt} and the subsidy rate $1 - p^h_{k_{jt}}$ in that year. I augment each observation x_{jt} with the extra information required to integrate p_k out by letting each observation be defined as $\tilde{x}_{jt} := (k_{jt}, p_{k_{jt}}, \omega_{jt}, i_{jt}, \tau_{jt}, p^h_{k_{jt}}, d^o_{jt})$, where d^o_{jt} is a dummy variable taking value 1 for the firm-year observations that need to be integrated out. For completeness, let $\tau_{jt} = -\infty, p^h_{k_{jt}} = 1$ when $d^o_{jt} = 0$.

For the observations where p_k is observed the log likelihood is $\ell^o(\tilde{x}_{jt}|\theta) = \ln L(k_{jt}, p_{k_{jt}}, \omega_{jt}, i_{jt}|\theta)$. Conversely, for the observations where p_k is unobserved p_k is equal to $p^h_{k_{jt}}$ with probability $\lambda^{t-\tau_{jt}}$ and is equal to 1 with probability $1-\lambda^{t-\tau_{jt}}$. For these observations the log likelihood is given by

$$\ell^u(\tilde{x}_{jt}|\theta) = \ln \Big[(1 - \lambda_k^{t-\tau_t})L(k_{jt}, 1, \omega_{jt}, i_{jt}|\theta) + \lambda_k^{t-\tau_t} L(k_{jt}, p^h_{k_{jt}}, \omega_{jt}, i_{jt}|\theta)\Big]$$

Consequently, the log likelihood of the data is

$$\mathcal{L}(\{\tilde{x}_{jt}\}_{jt}; \theta) = \sum_{j=1}^{N} \sum_{t=1}^{T} [\ell^o(\tilde{x}_{jt}|\theta)]^{d^o_{jt}} [\ell^u(\tilde{x}_{jt}|\theta)]^{1-d^o_{jt}}$$

$$d^o_{jt} = \begin{cases} 1 & \text{if } p_{k_{jt}} \text{ is observed} \\ 0 & \text{otherwise} \end{cases}$$

8.5.2 Maximum likelihood estimator

The solution of the maximum likelihood problem provides an estimate of the adjustment cost parameter c_2, the persistence of the subsidy policy rule λ_k and the variance σ^2_ϵ of the structural error. Specifically, the maximum likelihood problem is defined as

$$\begin{aligned}
&\max_{\theta} && \mathcal{L}(\{\tilde{x}_{jt}\}_{it};\theta) \\
&\text{subject to:} && V(\omega,p_k,k,\epsilon) = \max_{i\geq 0} \pi(\omega,k) - \left[ip_k + \frac{c_2}{2}\frac{i^2}{k}\right]\epsilon + \\
& && +\beta\int_{\omega',p_k',\epsilon'} V(\omega',p_k',k',\epsilon')dF(\omega',p_k'|\omega,p_k)dF_\epsilon(\epsilon') \\
& && F(\omega',p_k'|s) = F_\omega(\omega'|\omega)F_{p_k}(p_k'|p_k,\omega) \\
& && k' = i + (1-\delta)k,\ c_2 > 0
\end{aligned} \tag{33}$$

This is a constrained maximization problem in which the constraint requires that firms behave optimally. The productivity process F_ω is estimated in section 6 and is considered known and the $F_{p_k}^U$ component of F_{p_k} is already estimated in section 8.4. Given parameter values $\theta = (c_2, \lambda_k, \sigma_\epsilon^2)$ the Bellman equation in (33) defines a unique V and therefore a unique EV. Formally, the Bellman equation defines a mapping $\theta \mapsto V$ and a mapping $T\Theta : \theta \mapsto EV$. If the mapping $T\Theta$ has a closed-form representation, the maximum likelihood problem reduces to an unconstrained maximum likelihood problem. Unfortunately, $T\Theta$ does not have a closed-form representation and for each θ, $EV = T\Theta(\theta)$ needs to be approximated numerically. Fernández-Villaverde, Rubio-Ramírez and Santos (2006) show that, under regularity conditions, in maximum likelihood estimation problems where the value function and the optimal policy function are numerically calculated, as in (33), the approximated log likelihood converges to the true log likelihood as the approximation error in the calculation of the value function goes to zero. Hence, inference based on the solution of (33) is valid.

Since investment is a dynamic choice expectations over future states are important. The dependence of $F_{p_k}(p_k'|p_k = 1, \omega)$ on ω allows for heterogeneity in firm's expectations stemming from the self-selection of firms into the policy. Only productive firms have high probability to be subsidized in the future both because they are the ones that will apply for a subsidy, as they are the ones that want to invest, and possibly because the government has a preference to allocate subsidy grants to the most productive firms.

8.6 Numerical implementation of the MLE

Since there is no closed form representation of $T\Theta : \theta \mapsto EV$, for each θ, $EV = T\Theta(\theta)$ needs to be approximated numerically. There are three main decisions to be made when attempting to solve the problem in (33) numerically. The first is how to approximate EV in some finite space, the second is how to solve the functional equation in (33) for EV and the third is which optimization algorithm to use.

Note that in order to calculate the likelihood the econometrician needs to approximate EV which does not necessarily require the approximation of V itself. Using the insight of Rust (1987) for discrete choice dynamic models and its extension by Fuentes, Gilchrist and Rysman (2006) I can reduce the dimensionality of the approximation space by solving a functional equation in EV instead of V.

Integrating both sides of the Bellman equation (28) and using the law of iterated expectations (L.I.E.) and the serial independence of ϵ yields

$$\int_\epsilon V(\omega, p_k, k, \epsilon) = \int_\epsilon \Bigg\{ \max_{k' \geq (1-\delta)k} \pi(\omega, k) - \left[ip_k + C(k, k')\right] \epsilon$$
$$+\beta \int_{\omega', p_k', \epsilon'} V(\omega', p_k', k', \epsilon') dF(\omega', p_k' | \omega, p_k) dF_\epsilon(\epsilon') \Bigg\} dF_\epsilon(\epsilon)$$
$$\Leftrightarrow \ \mathcal{E}V(\omega, p_k, k) = \int_\epsilon \Bigg\{ \max_{k' \geq (1-\delta)k} \pi(\omega, k) - \left[ip_k + C(k, k')\right] \epsilon +$$
$$+\beta \int_{\omega', p_k'} \mathcal{E}V(\omega', p_k', k') dF(\omega', p_k' | \omega, p_k) \Bigg\} dF_\epsilon(\epsilon),$$
$$\mathcal{E}V(\omega, p_k, k) := \int_\epsilon V(\omega, p_k, k, \epsilon) dF_\epsilon(\epsilon)$$
$$\int_{\omega', p_k', \epsilon'} V(\omega', p_k', k', \epsilon') dF(\omega', p_k' | \omega, p_k) dF_\epsilon(\epsilon') \overset{\text{L.I.E.}}{=}$$
$$\int_{\omega', p_k'} \underbrace{\left[\int_{\epsilon'} V(\omega', p_k', k', \epsilon') dF_\epsilon(\epsilon')\right]}_{\mathcal{E}V(\omega', p_k', k')} dF(\omega', p_k' | \omega, p_k)$$

To calculate each firm's optimal investment behavior knowledge of $\mathcal{E}V$ is sufficient and I exploit this fact to reduce the computational burden of the estimation . Moreover, the mapping

$$\mathcal{E}V(\omega, p_k, k) \mapsto \int_\epsilon \Bigg\{ \max_{k' \geq (1-\delta)k} \pi(\omega, k) - \left[ip_k + C(k, k')\right] \epsilon + \beta \int_{\omega', p_k'} \mathcal{E}V(\omega', p_k', k') dF(\omega', p_k' | \omega, p_k) \Bigg\}$$

satisfies Blackwell's monotonicity and discounting conditions which are sufficient for a contraction mapping. As a result, $\mathcal{E}V$ can be calculated with value function iteration. Let the current state vector be $s = (\omega, p_k, k)$ and the next period state vector be $s' = (\omega', p_k', k')$, then the fixed point equation in $\mathcal{E}V$ is

$$\begin{aligned} \mathcal{E}V(s) = \int_{\epsilon} \left\{ \max_{i \geq 0} \pi(\omega, k) + \left[ip_k + C(k,k')\right]\epsilon + \beta E_F \mathcal{E}V(s') \right\} dF_{\epsilon}(\epsilon) \\ F(\omega', p_k' | \omega, p_k) = F_{\omega}(\omega'|\omega) F_{p_k}(p_k'|p_k, \omega) \\ k' = i + (1-\delta)k, \; C(k,k') > 0 \end{aligned} \tag{34}$$

The approximation scheme of the $\mathcal{E}V$ is dictated by the calculation of ϵ in (31) which involves taking the derivative of $\mathcal{E}V$ at every data-point meaning that $\mathcal{E}V$ needs to be approximated in a space of differentiable functions. I choose to approximate $\mathcal{E}V$ with Chebyshev polynomials which are known to have very desirable approximation properties.

There are mainly two ways to solve the functional equation after the choice of the function approximation space. The first method is to use value function iteration[25] and the second is to use a non-linear equation-solving software to find the Chebyshev coefficients that satisfy (34). One can utilize both of these solution methods in different stages of the numerical implementation. As for the optimization algorithm, one can use a combination of a grid-search algorithm and an non-linear optimization algorithm that uses derivative information. Specifically, it is possible to use a grid search and VFI first to picture of how the likelihood is shaped over a grid of the parameter space and then use the maximand of the likelihood found with the grid search algorithm as a starting value for the non-linear optimizer which, in turn, finds the optimum with accuracy. The advantage of the non-linear solver is that it uses derivative information to find the optimum accurately but its disadvantage is that it might converge very slowly if the starting value is very distant from the optimum. In this current version of the paper I use only grid search with value function iteration to calculate the estimates. For details on the computation of the value function and the numerical implementation of the maximum likelihood problem see appendix D.

[25]Recall that in section 4.3.2 I showed that (34) defines a contraction mapping in $\mathcal{E}V$ so value function iteration (VFI) is one of the options to solve for $\mathcal{E}V$.

8.7 Dynamic parameter estimates

The estimates of the parameter c_2, which defines the capital adjustment cost function, the persistence in the subsidy policy rule λ_k and the variance σ_ϵ^2 of the distribution of the idiosyncratic *iid* shock ϵ are reported in this section.

The results of the maximum likelihood estimation algorithm are presented in table 20. The adjustment cost parameter is $2e - 7$ which means that the adjustment cost of the median firm in the sample is less than .01 percent of its investment expenditure. This estimate implies that adjustment costs are economically insignificant. While my adjustment cost estimate is not directly comparable with estimates from other datasets due to differences in modeling and sample selection criteria, it is useful to report some estimates from the empirical literature on investment models. Cooper and Haltiwanger (2006) using panel data for US manufacturing plants in a model with complete investment reversibility find a c_2 of .45 while Bloom (2009) in a model similar to Cooper and Haltiwanger (2006) estimates c_2 to be 9.6. using a sample of US publicly traded firms. Asker, Collard-Wexler and De Loecker (2014) using panel data from many countries and a model with quadratic and fixed adjustment costs estimate c_2 to be from .42 for France to 17.6 for the USA. These papers all use simulated method of moments to estimate the structural parameters. Fuentes, Gilchrist and Rysman (2006) using data from Chilean Manufacturing plants estimate c_2 to be 8e-4 and they also estimate the degree of investment irreversibility to be close to complete irreversibility. To estimate their parameters they use a maximum likelihood estimator similar to the estimator developed in this paper. My results are more in line with Fuentes, Gilchrist and Rysman (2006) which is not surprising since I assume complete irreversibility and I have the same estimation methodology. My estimate of the adjustment cost parameter suggests that, in the presence of complete or almost-complete irriversibility, adjustment cost is not needed to justify firm's observed investment behavior.

The persistence in the subsidy policy λ_k is estimated to be .34 which means that a subsidized firm will receive the same subsidy rate next period with probability .34 and will receive no subsidy with probability .66. This parameter implies that the probability for a firm to receive a subsidy for three consecutive years is low and equal to 12 percent, which is consistent with the fact that both the investment rate and the investment level drop more than 50 percent three years after firms are granted the subsidy.

The standard deviation of $\ln(\epsilon)$ is $\sigma_\epsilon = .2$ which implies that a firm with a shock ϵ at the 90th percentile of the structural error distribution faces cost of investment 30% higher than a firm with a shock equal to the median of the distribution.

8.8 Counterfactual analysis

The estimation of the policy and adjustment cost parameters is an essential step to quantify the effects of the subsidy policy on firms' behavior and aggregate investment. In this section I use the estimated parameters of the model to investigate how the subsidized and unsubsidized firms' behavior is affected by the policy and measure the level of aggregate investment that would have been realized in the absence of the policy. I begin by stating the underlying assumptions necessary to support the quantitative results of this section.

The partial equilibrium model presented in this paper is developed under the assumption that any change in the investment subsidy policy has no general equilibrium effects, namely factor prices remain unchanged. A sufficient condition for this to happen is that the demand for materials, electricity, labor, machinery and transportation equipment, and land coming from the food manufacturing sector is small relatively to the aggregate demand in these factors markets. Since Greece is a small open economy factor prices of machinery, transportation equipment, electricity, and materials (such as agricultural products) are determined in international markets so that assuming no general equilibrium effects is plausible. As for labor, it is sufficient to assume that labor supply in not specific for food manufacturing, not spatially segmented, and the skill composition of workers is homogeneous across manufacturing sectors. As for land, it is reasonable to assume that manufacturing is not a land intensive sector so that it is unlikely for any investment subsidy policy to affect land prices.

In theory, the subsidy policy should affect subsidized firms directly by altering their price of capital and unsubsidized firms indirectly by affecting their option value of waiting to invest in a future period when there may be a chance to be subsidized. Therefore, the effects on both subsidized and unsubsidized firms' behavior need to be analyzed to assess the economic significance and effectiveness of the policy. Figures 6 and 7 depict the heterogeneous effect of the subsidy on the optimal firm investment implied by the estimated parameters of the model. Figure 6 depicts the optimal investment for a firm at the 90th percentile of the productivity distribution. The dashed line shows the optimal investment of a firm subsidized with a 50 percent subsidy rate. The continuous line depicts the optimal investment of an unsubsidized firm while the continuous dotted line depicts the optimal investment of a firm in an environment without a subsidy policy. Note that subsidized firms invest more than unsubsidized firms and unsubsidized firms invest less than what they would invest in an environment without the subsidy policy. In the terminology of section 7, the expected difference between the dashed and the continuous line is the average treatment effect (ATE). In the terminology of section 7.1 the expected difference between the dashed and the dotted line is the policy relevant treatment effect on the treated (PRTET) while the expected difference between the continuous and the dotted line is the policy relevant treatment effect on the untreated (PRTEU). The negative sign of PRTEU implied by the graph is consistent with the reduced form evidence of section 7.1 indicating that unsubsidized firms with high probability of being subsidized in the future invest less than they would in an environment without the subsidy policy. Figure 7 depicts the optimal investment for a firm at the 50th percentile of the productivity distribution. Qualitatively, the optimal investment policy is the same for a firm at the 90th percentile and at the 50th percentile of the productivity distribution. However, the PRTEU effect for low productivity firms is economically insignificant. This result is also in line with the reduced form evidence of section 7.1 indicating that unsubsidized firms with low expectations of being subsidized in the future do not change their investment behavior because of the policy. This heterogeneous effect of the policy on unsubsidized firms highlights the importance of taking into account unobserved heterogeneity.

To study the aggregate implications of the subsidy policy I condition on the productivity realizations and the initial capital stock for each firm in the sample. I perform the counterfactual experiments conditioning on productivity because the productivity process is exogenous in my model and hence, policy invariant. To evaluate the expected effect of the no policy case I simulate two economies which have the same productivity realizations, and share the same Markov process F_ω, profit function π, and adjustment cost parameter c_2. The only difference between these two economic environments is the presence of the estimated subsidy policy $\hat{F}_{p_k}$ in the first and the absence of it in the second. My results show that in the absence of the policy, aggregate investment would have been 11 percent higher. This is a surprising result because it suggests that the policy actually decreases investment in the population instead of boosting it. Subsidized firms invest more but unsubsidized firm invest less due to the policy because the policy increases their option value of waiting. This result could imply that subsidizing a small number of firms with large subsidy rates might in fact decrease the aggregate investment because the effect of the policy on unsubsidized firms is large enough to overpower the effect on subsidized firms.

9 Conclusions

This paper develops and estimates a dynamic investment model of heterogeneous firms to evaluate the effect of a Greek investment subsidy policy on the sectoral aggregate investment in the food and beverages manufacturing sector. I find that there is significant heterogeneity in productivity among firms and that firms self-select into the policy according to their productivity. More specifically, highly productive firms have much higher probability of receiving a subsidy than lower productivity ones and I control for this selection in the estimation of the model parameters. The policy induces subsidized firms to invest more while it induces highly productive unsubsidized firms that have a high probability to be subsidized in the future to invest less because the policy increases their option value of waiting. Overall, sectoral aggregate investment decreases by 11 percent.

The analysis carried out in the paper suggests promising avenues for future research. A deeper investigation of the scope and the extent of the subsidy policy involves exploring how the regional character of the investment subsidy program affects the allocation of capital across regions and how a location-independent subsidy policy could have a different impact on aggregate investment and productivity. To analyze these issues requires to re-parametrize the subsidy policy rule making it location-dependent and include a location variable to the state vector in the firm's dynamic programming problem. As the data explicitly include a location variable, the estimation of the re-parametrized model is feasible and is the next step on my research agenda.

References

Adda, Jérôme and Russell Cooper (2000) "Balladurette and Juppette: A Discrete Analysis of Scrapping Subsidies," *Journal of Political Economy*, Vol. 108, pp. 778–806. [back to pg. 8]

——— (2003) *Dynamic Economics: Quantitative Methods and Applications*, Cambridge, MA: MIT press. [back to pg. 46]

Angelucci, Manuela and Giacomo De Giorgi (2009) "Indirect Effects of an Aid Program: How Do Cash Transfers Affect Ineligibles' Consumption?" *American Economic Review*, Vol. 91, pp. 486–508. [back to pg. 8]

Asker, John, Allan Collard-Wexler, and Jan De Loecker (2014) "Dynamic Inputs and Resource (Mis)Allocation," *Journal of Political Economy*, Vol. 122, pp. 1013–1063. [back to pg. 2], [back to pg. 7], [back to pg. 10], [back to pg. 13], [back to pg. 42], [back to pg. 46], [back to pg. 53]

Aw, Bee Y., Mark J. Roberts, and Daniel Y. Xu (2011) "R&D Investment, Exporting, and Productivity Dynamics," *American Economic Review*, Vol. 101, pp. 1312–1344. [back to pg. 10], [back to pg. 12], [back to pg. 25], [back to pg. 26]

Bloom, Nicholas (2009) "The Impact of Uncertainty Shocks," *Econometrica*, Vol. 77, pp. 623–685. [back to pg. 2], [back to pg. 7], [back to pg. 53]

Bond, Stephen and John Van Reenen (2007) "Microeconometric Models of Investment and Employment," in J. J. Heckman and E. E. Leamer eds. *Handbook of Econometrics*, Vol. 6A: Elsevier, pp. 4417–4498. [back to pg. 46]

Caballero, Ricardo J. and Eduardo M. R. A. Engel (1999) "Explaining Investment Dynamics in U.S. Manufacturing: A Generalized (S,s) Approach," *Econometrica*, Vol. 67, pp. 783–826. [back to pg. 6]

Cai, Yongyang and Kenneth L. Judd (2010) "Stable and Efficient Compuational Methods for Dynamic Programming," *Journal of the European Economic Association*, Vol. 8, pp. 626–634. [back to pg. 84]

——— (2013) "Shape Preserving Dynamic Programming," *Mathematics of Operations Research*, Vol. 77, pp. 407–421. [back to pg. 84]

——— (2014) "Advances in Numerical Dynamic Programming and New Applications," in Karl Schmedders and Kenneth L. Judd eds. *Handbook of Computational Economics*, Vol. 3: Elsevier, pp. 479–516. [back to pg. 84]

Cai, Yongyang, Kenneth L. Judd, Thomas S. Lontzek, Valentina Michelangeli, and Che-Lin Su (2013) "Nonlinear Programming Method for Dynamic Programming," *NBER wp*. [back to pg. 84], [back to pg. 86]

Cooper, Russell and John C. Haltiwanger (2006) "On the Nature of Capital Adjustment Costs," *Review of Economic Studies*, Vol. 73, pp. 611–633. [back to pg. 2], [back to pg. 6], [back to pg. 7], [back to pg. 10], [back to pg. 33], [back to pg. 42], [back to pg. 46], [back to pg. 53]

Criscuolo, Chiara, Ralf Martin, Henry Overman, and John Van Reenen (2012) "The Causal Effects of an Industrial Policy," *NBER Working Paper*, Vol. 17842. [back to pg. 7]

Dixit, Avinash K. and Robert S. Pindyck (1994) *Investment under Uncertainty*, Princeton, NJ: Princeton University Press. [back to pg. 6], [back to pg. 19]

Doms, Mark and Timothy Dunne (1998) "Capital Adjustment Patterns in Manufacturing Plants," *Review of Economic Dynamics*, Vol. 1, pp. 409–429. [back to pg. 6]

Doraszelski, Ulrich and Jordi Jaumandreu (2013) "R&D and Productivity: Estimating Endogenous Productivity," *Review of Economic Studies*, Vol. 80, pp. 1338–1383. [back to pg. 12]

Fazzari, Steven M., R. Glenn Hubbard, and Bruce C. Petersen (2000) "Investment-Cash Flow Sensitivities are Useful: A Comment on Kaplan and Zingales," *Quarterly Journal of Economics*, Vol. 115, pp. 695–705. [back to pg. 6]

Fernández-Villaverde, Jesús, Juan F. Rubio-Ramírez, and Manuel S. Santos (2006) "Convergence Properties of the Likelihood of Computed Dynamic Models," *Econometrica*, Vol. 74, pp. 93–119. [back to pg. 49]

Fisher, Irving (1907) *The Rate of Interest: Its nature, determination and relation to economic phenomena*, New York: Macmillan. [back to pg. 6]

Fuentes, Olga, Simon Gilchrist, and Marc Rysman (2006) "Irreversibility and Investment Dynamics for Chilean Manufacturing Plants: A Maximum Likelihood Approach," *Boston University Working Paper*. [back to pg. 2], [back to pg. 13], [back to pg. 42], [back to pg. 50], [back to pg. 53]

Gilchrist, Simon and Charles P. Himmelberg (1995) "Evidence on the Role of Cash Flow for Investment," *Journal of Monetary Economics*, Vol. 36, pp. 541–572. [back to pg. 6], [back to pg. 42]

Gourieroux, Christian S., Alain Monfort, and Eric M. Renault (1993) "Indirect Inference," *Journal of Applied Econometrics*, Vol. 8, pp. 85–118. [back to pg. 46]

Hansen, Lars Peter and Kenneth J. Singleton (1982) "Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models," *Econometrica*, Vol. 50, pp. 1269–1286. [back to pg. 45]

Hayashi, Fumio (1982) "Tobin's Marginal Q and Average Q: A Neoclassical Interpretation," *Econometrica*, Vol. 50, pp. 213–224. [back to pg. 6], [back to pg. 46]

Heckman, J. J. and E. J. Vytlacil (2007) "Econometric Evaluation of Social Programs, Part I: Causal Models, Structural Models and Econometric Policy Evaluation," in J. J. Heckman and E. E. Leamer eds. *Handbook of Econometrics*, Vol. 6B: Elsevier, pp. 4779–4874. [back to pg. 34]

House, Christopher L. and Matthew D. Shapiro (2008) "Temporary Investment Tax Incentives: Theory with Evidence from Bonus Depreciation," *American Economic Review*, Vol. 98, pp. 737–768. [back to pg. 8]

Hsieh, Chang-Tai and Peter J. Klenow (2009) "Misallocation and Manufacturing TFP in China and India," *Quarterly Journal of Economics*, Vol. 124, pp. 1403–1448. [back to pg. 2], [back to pg. 7]

Hubbard, R. Glenn, Anil K. Kashyap, and Toni M. Whited (1995) "Internal Finance and Firm Investment," *Journal of Money, Credit and Banking*, Vol. 27, pp. 683–701. [back to pg. 6]

Judd, Kenneth L. (1998) *Numerical Methods in Economics*, Cambridge, MA: MIT press. [back to pg. 82], [back to pg. 83]

Khan, Aubhik and Julia K. Thomas (2008) "Idiosyncratic Shocks and the Role of Nonconvexities in Plant and Aggregate Investment Dynamics," *Econometrica*, Vol. 76, pp. 395–436. [back to pg. 2]

Levinsohn, James and Amil Petrin (2003) "Estimating Production Functions Using Inputs to Control for Unobservables," *Review of Economic Studies*, Vol. 70, pp. 317–341. [back to pg. 25], [back to pg. 26]

Mian, Atif and Amir Sufi (2012) "The Effects of Fiscal Stimulus: Evidence from the 2009 Cash for Clunkers Program," *Quarterly Journal of Economics*, Vol. 127, pp. 1107–1142. [back to pg. 8]

Midrigan, Virgiliu and Daniel Yi Xu (2014) "Finance and Misallocation: Evidence from Plant-Level Data," *American Economic Review*, Vol. 104, pp. 422–458. [back to pg. 2], [back to pg. 7]

Modigliani, Franco and Merton H. Miller (1958) "The Cost of Capital, Corporation Finance and the Theory of Investment," *American Economic Review*, Vol. 48, pp. 261–297. [back to pg. 6]

Olley, G. Steven and Ariel Pakes (1996) "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, Vol. 64, pp. 1263–1297. [back to pg. 25]

Pakes, Ariel (1994) "Dynamic Structural Models, Problems and Prospects: Mixed Coninuous Discrete Controls and Market Interactions," in Christopher A. Sims ed. *Advances in Econometrics, Sixth World Congress*: Cambridge University Press, pp. 171–259. [back to pg. 46]

Peck, Stephen C. (1974) "Alternative Investment Models for Firms in the Electric Utilities Industry," *The Bell Journal of Economics and Management Science*, Vol. 5, pp. 420–458. [back to pg. 6]

Ralph, Daniel (2008) "Mathematical Programs with Complementarity Constraints in Traffic and Telecommunications Networks," *Philosophical Transactions of the Royal Society A*, Vol. 336, pp. 1973–1987. [back to pg. 87]

Restuccia, Diego and Richard Rogerson (2008) "Policy Distortions and Aggregate Productivity with Heterogeneous Establishments," *Review of Economic Dynamics*, Vol. 11, pp. 707–720. [back to pg. 2], [back to pg. 7]

Rincón-Zapatero, Juan Pablo and Manuel S. Santos (2009) "Differentiability of the Value Function without Interiority Assumptions," *Journal of Economic Theory*, Vol. 144, pp. 1948–1964. [back to pg. 46]

Rust, John (1987) "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," *Econometrica*, Vol. 55, pp. 999–1033. [back to pg. 50], [back to pg. 85]

——— (1988) "Maximum Likelihood Estimation of Discrete Control Processes," *SIAM Journal on Control and Optimization, 26, 1006-1024*, Vol. 26, pp. 1006–1024. [back to pg. 85]

Su, Che-Lin and Kenneth L. Judd (2012) "Constrained Optimization Approaches to Estimation of Structural Models," *Econometrica*, Vol. 80, pp. 2213–2230. [back to pg. 85]

Yagan, Danny (2014) "Capital Tax Reform and the Real Economy: The Effects of the 2003 Dividend Tax Cut," Unpublished Manuscript, UC Berkeley. [back to pg. 8]

A Data

For the construction of the capital stock series I combine information from the ICAP dataset and the Annual Survey of Manufactures (ASM). In the ICAP dataset the capital stock is broken down into four asset classes: Land, Machinery, Intangibles, and Sturctures, Equipment and Tranport equipment (see table 21). For each asset class there are two variables: acquisition value and current accumulated depreciation. I subtract accumulated depreciation from the acquisition value to obtain the book value of capital which is the closest measure to the value of capital in current prices[26]. $\tilde{A}_t^{rI}$ represents the value of capital of class rI at time t in current prices.

In the ASM dataset the gross investment is broken down into five asset classes: Land, Structures, Transport Equipment, Machinery, and Intangibles and Remaining Equipment (see table 21). $\tilde{A}_t^{rA}$ represents the value of capital of class rA at time t in current prices and $\tilde{I}_t^{rA}$ is the gross investment of class rA at time t in current prices.

Investment flows and initial capital stock are deflated by GDP deflators from the Eurostat AMECO database[27] to create variables in constant prices $A_t^{rI}, A_t^{rA}, I_t^{rA}$.

The capital stock series of asset class r is created with the perpetual inventory method assuming that investment takes one year to become operational, i.e. $A_t^r = (1 - \delta_r)A_{t-1}^r + I_{t-1}^r$. Depreciation rates are industry specific and come from the EU KLEMS database which is based on the BEA depreciation rates. The specific values of depreciation rates and deflators used for each asset class is reported in table 21.

I use the ICAP dataset to obtain the capital stock for the first year in which data on investment flows from ASM are available. Then I combine the initial capital stock from ICAP with the ASM investment flows to create the capital stock for each t. Unfortunately,

[26] The calculated value differs from the value of capital in current prices because it comprises different vintages acquired at different purchase prices.

[27] The GDP deflators provided in the AMECO database do not vary across industries.

the asset classes in the ICAP dataset do not correspond exactly to the asset classes in the ASM database. For this reason, in an extent to maintain as much consistency as possible, I create the capital stock as follows. Let t_0 be the year when the series begins in the ASM database, then

$$\begin{aligned} A_t^{rI} &= (1-\delta_{rI})^{t-t_0} A_{t_0}^{rI},\ \forall t_0 \leq t,\ \forall rI \\ A_t^{rA} &= (1-\delta_{rA}) A_{t-1}^{rA} + I_{t-1}^{rA},\ A_{t_0}^{rA} = 0,\ \forall t_0 < t,\ \forall rI \\ K_t &= \sum_{rA} A_t^{rA} + \sum_{rI} A_t^{rI},\ \forall t_0 \leq t \end{aligned}$$

Finally, I discard the observations with capital stock lest than the .5 percentile and more than the 99.5 percentile to avoid outliers with either too close to zero or too large values of capital.

B Formula Derivations

B.1 Underlying production function

Firm's production function is given by $q = \Omega k^{b_k} L^{b_L} M^{b_M}$, where Ω is total factor productivity, k is physical capital stock, L is labor, and M is materials input. Capital is quasi-fixed in the sense that a firm arrives at period t with capital stock k_t which will be used to produce period's t output q_t. This assumption is referred as time-to-build in the sense that any investment in physical capital at period t can be used in production at period $t+1$. The implication of this assumption for firm's investment decision is discussed in section 4.3.2. The firm buys labor and materials in competitive markets at prices w_M and w_L, rispectively and its total variable cost is TVC $= Mw_M + Lw_L$.

Assumption 1. *Constant returns in variable inputs:* $b_L + b_M = 1$

Assumption 1 implies that there are increasing returns to scale in production. Even though there are increasing returns to scale, the monopolistically competitive market structure and the consumer love for variety prevent the optimal scale of the firm to be unbounded. Firm's variable-cost minimization problem and cost function are given by

$$\begin{aligned}
\min_{L,k} Mw_M + Lw_L \text{ s. t. } q \leq \Omega k^{b_k} L^{b_L} M^{b_M} \Rightarrow \\
L = \frac{b_L w_M}{b_M w_L} M,\ M = \frac{q}{\Omega k^{b_k} B},\ B := \left[\frac{b_L w_M}{b_M w_L}\right]^{b_L} \\
TVC(w_L, w_M, q, k) = \frac{1}{\Omega k^{b_k}} w_L^{b_L} w_M^{b_M} \Gamma q,\ \Gamma = \left[\frac{b_L}{b_M}\right]^{b_L} + \left[\frac{b_M}{b_L}\right]^{b_L} \Rightarrow \\
MC(w_L, w_M, k) = \frac{\partial TVC}{\partial q} = \frac{1}{\Omega k^{b_k}} w_L^{b_L} w_M^{b_M} \Gamma \Rightarrow \\
\ln MC = -\ln \Omega - b_k \ln k + b_L \ln w_L + b_M \ln w_M + \ln \Gamma
\end{aligned} \tag{35}$$

The correspondence between equations (1) and (35) is $\ln \Omega \equiv \omega$, $-b_k \equiv \beta_k$, $\ln \Gamma \equiv \beta_0$, $(w_L, w_M) \equiv w$, $(b_L, b_M) \equiv \beta_w$, $\ln \Gamma \equiv \beta_0$.

B.2 Inverting the demand for materials with respect to producvitity

Given the demand for firm's output and firm's optimal price given in (2), output q and the log of the total material expenditure m are

$$
\begin{aligned}
q &= \left(\frac{\eta}{\eta+1} \frac{1}{\Omega k^{b_k}} w_L^{b_L} w_M^{b_M} \Gamma \right)^{\eta} \\
M = \frac{1}{\Omega^{\eta+1} k^{b_k(\eta+1)}} \left(\frac{\eta}{\eta+1} \right)^{\eta+1} \frac{\left(w_L^{b_L} w_M^{b_M} \Gamma \right)^{\eta}}{B} &= \frac{1}{\Omega^{\eta+1} k^{b_k(\eta+1)}} \Delta \Rightarrow \\
\text{TME} = w_M M &= w_M \frac{1}{\Omega^{\eta+1} k^{b_k(\eta+1)}} \Delta \Rightarrow \\
m = \ln \text{TME} &= \ln w_M + (\eta+1)\left(-\ln \Omega - b_k \ln k\right) + \ln \Delta
\end{aligned}
\tag{36}
$$

From (36), since by definition $\eta + 1 < 0$, m is a strictly increasing function of $\omega = \ln \Omega$. Consequently, there exists a function ω with $\omega = \omega(m, k)$ that is strictly increasing in m. Function ω is crucial in the estimation of the profit function in section 6. More specifically, function ω is the theoretical argument behind the use of materials expenditure m to control for unobserved productivity ω_{jt} in equation (10).

C Figures and Tables

Table 1: Summary statistics of the production data

Variable	Med	Mean	Std	Min	Max	%Zeros
Capital k	2.4	6.7	11	.01	88	0
Investment i	.23	1	2.7	0	73	6.3
Variable profit* π	1.3	5.4	13	.0007	254	0
Materials expense m	2.5	7.7	15	0	177	1.4
Electricity expense e	.07	.19	.38	0	4.6	.1
Labor cost lc	.83	2.8	5.7	0	70	.06
No. of employees	46	115	194	0	1942	.06
Sales† r	5.2	16	31	.3	396	0
No. of Firms	430					
Firm-Year Obs.	2867					
Years				1999	2009	
$corr(\ln(k), \ln(\pi))$	.75					

Monetary variables are expressed in million 2005 Euros. For details on the deflators see appendix A.
† Sales include revenue from the sale of manufactured goods and revenue from subcontracting but exclude revenue from purely trading activity.
* Variable profits are defined as revenue minus variable input cost, i.e. $\pi = r - m - e - lc$.

Table 2: Summary statistics of the subsidy data

Variable	25^{th}	Med	75^{th}	Mean	St.dev.	Min	Max	Row sum
Subsidized project size*	.68	1.5	2	2.1	2.9	.11	28	378
Cash subsidy rate†	.3	.4	.45	.39	.097	.057	.55	-
Grant Year	2005	2006	2008		1999	2009		
No. of Subsidies	183							
No. of Subsidized Firms††	134							

Monetary variables are expressed in million 2005 Euros. For details on the deflators see appendix A.
* Project size represents the maximum investment expenditure for which the subsidy rate applies.
† Subsidy rate is the fraction of the investment expenditure the firm receives in cash from the subsidy program.
†† The number of subsidized firms represents the number of firms that received at least one subsidy during the years 1999-2009. Each firm can be subsidized more than once during the 11 year period included in the sample, thus the number of subsidies granted is larger than the number of subsidized firms.

Table 3: Subsidized firms' aggregate investment

	INVESTMENT EXPENDITURE PATTERNS			
	Total Investment	% of Aggregate Investment	Obs.	% of all obs.
Obs. within 3 years after each subsidy decision* $0 \leq t - \tau_t \leq 2$	330	11	208	7
All the rest firm-years $t - \tau_t < 0$ or $2 < t - \tau_t$	2673	89	2659	93
Column sum	3004	100	2867	100

Monetary variables are expressed in million 2005 Euros. For details on the deflators see appendix A.
t is the current year and τ_t is the year the firm received the latest subsidy. For firms that have never received a subsidy until time t $\tau_t = -\infty$ and, consequently, $t - \tau_t = \infty$.
* These are firm-year observations corresponding to either the same year a firm was granted a subsidy or the two subsequent years.

Table 4: Subsidized firms' investment patterns

Years since the year t_0 the subsidy was granted		Median Investment	Median Inv. Rate	Obs.
Before the subsidy	$t < t_0$	0.34	0.13	632
Year of the subsidy grant	$t = t_0$	1.12	0.15	78
One year after the grant	$t - t_0 = 1$	0.92	0.14	71
Two years after the grant	$t - t_0 = 2$	0.65	0.11	64
Three or more years after	$t - t_0 \geq 3$	0.36	0.07	104

Monetary variables are expressed in million 2005 Euros.
t is the current year and t_0 is the year the decision to grant a firm a subsidy was taken.

Table 5: Comparison between subsidized and never-subsidized firms

	MEDIAN		St.dev.	Median	Mean	Observ-
	Investm.	Inv. Rate	Inv. Rate	Capital	Capital	ations
Subsidized firms before receipt of the subsidy	0.34	0.13	0.98	2.99	6.78	632
Firms never subsidized during the sample years	0.16	0.08	0.89	1.81	5.24	1918

Monetary variables are expressed in million 2005 Euros.

Table 6: Greek administrative regions categorization according to the subsidy intensity

Region	Binary Categorization
Attica	L
Central Greece	L
Central Macedonia	L
Crete	L
Epirus	H
East Macedonia and Thrace	H
Ionian Islands	L
North Aegean	H
Peloponese	H
South Aegean	L
Thessaly	L
West Macedonia	L
Western Greece	H

Table 7: Statistics by binary (H, L) location

Location	Median			Subsidy	%appli-	%subsi-	odds[‡]	mean
	k[§]	N_L[¶]	N_L/k	rate	cants[*]	dized[†]		wait[**]
L	2.4	46	21	.4	29	24	.83	195
H	2.7	45	18	.45	44	39	.88	253

[§] k=Capital stock expressed in million 2005 Euros.
[¶] N_L=Number of employees
[*] % of firms in the sample which applied for a subsidy at least once
[†] % of firms in the sample which were granted a subsidy at least once
[‡] Probability of receiving a subsidy conditional on applying for a subsidy
[**] Average number of days between an application and the result of the application process conditional on success.

Table 8: Reduced form log investment demand

Variable	Model (1)	Model (2)	Model(3)
Treatment D^2	$.6(.1)^*$	-	-
Treatment D^3	-	$.5(.1)^*$	-
Treatment D^4	-	-	$.5(.1)^*$
ω	$1.8(.19)^*$	$1.7(.19)^*$	$1.7(.19)^*$
ω^2	-.2(.2)	-.2(.2)	-.2(.2)
$\ln k$	$.7(.04)^*$	$.7(.04)^*$	$.7(.04)^*$
$(\ln k)^2$	-.02(.02)	-.02(.02)	-.02(.02)
$\omega \ln k$	-.06(.09)	-.06(.09)	-.07(.09)
Sample size	2651	2651	2651

Linear regression model with variables in logarithmic scale. Robust standard errors are reported.

Model (1): $\ln i = \gamma_{ATE} D^2_{jt} + \gamma_L D^l_{jt} + \sum_{t=1}^{T} \gamma_t d^t + \gamma_k \ln k_{jt} + \gamma_{kk} (\ln k_{jt})^2 + \gamma_\omega \omega_{jt} + \gamma_{\omega\omega} (\omega_{jt})^2 + \gamma_{k\omega} \omega_{jt} \ln k_{jt},\ i > 0$

Model (2): $\ln i = \gamma_{ATE} D^3_{jt} + \gamma_L D^l_{jt} + \sum_{t=1}^{T} \gamma_t d^t + \gamma_k \ln k_{jt} + \gamma_{kk} (\ln k_{jt})^2 + \gamma_\omega \omega_{jt} + \gamma_{\omega\omega} (\omega_{jt})^2 + \gamma_{k\omega} \omega_{jt} \ln k_{jt},\ i > 0$

Model (3): $\ln i = \gamma_{ATE} D^4_{jt} + \gamma_L D^l_{jt} + \sum_{t=1}^{T} \gamma_t d^t + \gamma_k \ln k_{jt} + \gamma_{kk} (\ln k_{jt})^2 + \gamma_\omega \omega_{jt} + \gamma_{\omega\omega} (\omega_{jt})^2 + \gamma_{k\omega} \omega_{jt} \ln k_{jt},\ i > 0$

Table 9: Reduced form investment demand

Variable	Model (1)	Model (2)	Model(3)
Treatment D^2	.4(.16)*	-	-
Treatment D^3	-	.34(.14)*	-
Treatment D^4	-	-	.26(.13)*
$\exp(\omega)$	1(.18)*	1(.18)*	1(.18)*
k	.07(.03)*	.07(.03)*	.07(.03)*
$\exp(\omega)k$	-2e-5(8e-3)	-9e-5(8e-3)	-6e-5(8e-3)
Sample size	2828	2828	

Tobit regression with investment i as the dependent variable. Both investment and capital are expressed in million 2005 Euros. Robust standard errors are reported.
Model (1): $i = \gamma_{ATE} D^2_{jt} + \gamma_L D^l_{jt} + \sum_{t=1}^{T} \gamma_t d^t + \gamma_k k_{jt} + \gamma_\omega \exp(\omega_{jt}) + \gamma_{k\omega} \exp(\omega_{jt}) k_{jt}$

Model (2): $i = \gamma_{ATE} D^3_{jt} + \gamma_L D^l_{jt} + \sum_{t=1}^{T} \gamma_t d^t + \gamma_k k_{jt} + \gamma_\omega \exp(\omega_{jt}) + \gamma_{k\omega} \exp(\omega_{jt}) k_{jt}$

Model (3): $i = \gamma_{ATE} D^4_{jt} + \gamma_L D^l_{jt} + \sum_{t=1}^{T} \gamma_t d^t + \gamma_k k_{jt} + \gamma_\omega \exp(\omega_{jt}) + \gamma_{k\omega} \exp(\omega_{jt}) k_{jt}$

Table 10: Reduced form investment rate

Variable	Model (1)	Model (2)	Model(3)
Treatment D^2	.09(.04)*	-	-
Treatment D^3	-	.09(.04)*	-
Treatment D^4	-	-	.09(.04)*
ω	.9(.2)*	.9(.2)*	.9(.2)*
k	-.16(.04)*	-.16(.04)*	-.16(.04)*
$\exp(\omega)k$	-.14(.04)*	-.14(.04)*	-.14(.04)*
Sample size	2828	2828	2828

Tobit regression with investment rate i/k as the dependent variable. Both investment and capital are expressed in million 2005 Euros. Robust standard errors are reported.
Model (1): $i/k = \gamma_{ATE} D^{sp}_{jt} + \gamma_L D^l_{jt} + \sum_{t=1}^{T} \gamma_t d^t + \gamma_k \ln k_{jt} + \gamma_\omega \omega_{jt} + \gamma_{k\omega} \omega_{jt} \ln k_{jt}$

Model (2): $i = \gamma_{ATE} D^3_{jt} + \gamma_L D^l_{jt} + \sum_{t=1}^{T} \gamma_t d^t + \gamma_k k_{jt} + \gamma_\omega \exp(\omega_{jt}) + \gamma_{k\omega} \exp(\omega_{jt}) k_{jt}$

Model (3): $i = \gamma_{ATE} D^4_{jt} + \gamma_L D^l_{jt} + \sum_{t=1}^{T} \gamma_t d^t + \gamma_k k_{jt} + \gamma_\omega \exp(\omega_{jt}) + \gamma_{k\omega} \exp(\omega_{jt}) k_{jt}$

Table 11: Investment spike statistics

Spike Dummy D^{spike}	% of obs.	Aggregate inv. share
0	74	.38
1	26	.62

Table 12: Spike probit

Variable	Model (1)	Model (2)	Model(3)
Treatment D^2	.4(.1)*	-	-
Treatment D^3	-	.36(.1)*	-
Treatment D^4	-	-	.38(.1)*
ω	1(.16)*	1(.16)*	1(.16)*
ω^2	-.02(.2)	-.005(.2)	.01(.2)
$\ln k$	-.17(.04)*	-.17(.04)*	-.17(.04)*
$(\ln k)^2$	-.02(.01)	-.02(.01)	-.02(.02)
$\omega \ln k$	-.08(.08)	-.08(.08)	-.08(.08)
Sample size	2828	2828	2828

Probit model with D^{spike} as the dependent variable. Robust standard errors are reported.
Model (1): $\text{Prob}(D^{\text{spike}} = 1|\cdot) = \Phi[\gamma_{ATE} D^2_{jt} + \gamma_L D^l_{jt} + \sum_{t=1}^{T} \gamma_t d^t + \gamma_k \ln k_{jt} + \gamma_{kk} (\ln k_{jt})^2 + +\gamma_\omega \omega_{jt} + \gamma_{\omega\omega} (\omega_{jt})^2 + \gamma_{k\omega} \omega_{jt} \ln k_{jt}]$

Model (2): $\text{Prob}(D^{\text{spike}} = 1|\cdot) = \Phi[\gamma_{ATE} D^3_{jt} + \gamma_L D^l_{jt} + \sum_{t=1}^{T} \gamma_t d^t + \gamma_k \ln k_{jt} + \gamma_{kk} (\ln k_{jt})^2 + +\gamma_\omega \omega_{jt} + \gamma_{\omega\omega} (\omega_{jt})^2 + \gamma_{k\omega} \omega_{jt} \ln k_{jt}]$

Model (3): $\text{Prob}(D^{\text{spike}} = 1|\cdot) = \Phi[\gamma_{ATE} D^4_{jt} + \gamma_L D^l_{jt} + \sum_{t=1}^{T} \gamma_t d^t + \gamma_k \ln k_{jt} + \gamma_{kk} (\ln k_{jt})^2 + +\gamma_\omega \omega_{jt} + \gamma_{\omega\omega} (\omega_{jt})^2 + \gamma_{k\omega} \omega_{jt} \ln k_{jt}]$

Table 13: Statistics before and after the policy change

Year t	%Total Cash Allocated†	Median Subs. Rate	mean wait	% of grants
$1999 \leq t \leq 2003$	8	.3	120	15
$t \geq 2004$	92	.4	200	85

† in millions 2005 Euros

Table 14: Diff-in-diff reduced form log investment for unsubsidized firms

Variable	Model (1)	Model (2)	Model(3)
Dif-in-Dif: $D^{after}D^{l}$	-.7(.2)*	.16(.16)	-.14(.12)
Periods after change: D^{after}	.05(.1)	-.3(.1)*	-.23(.07)*
Location H: D^{l}	.03(.1)	-.3(.1)*	-.08(.08)
ω	1.2(.3)*	1.8(.2)*	1.7(.2)*
ω^2	.7(.4)	-.4(.25)	-.2(.2)
$\ln k$	.7(.1)*	.6(.05)*	.7(.04)*
$(\ln k)^2$	.01(.02)	-.03(.02)	-.02(.02)
$\omega \ln k$	-.45(.18)*	.07(.1)	-.08(.09)
Sample size	772	1573	2345
Share of investment	47%	53%	100%
Share of sales	39%	61%	100%

In all three models the same equation is estimated for different samples. Robust standard errors are reported.
$\ln i = \gamma_{DID} D^{after}_{jt} D^{l}_{jt} + \gamma_{after} D^{after}_{jt} + \gamma_L D^{l}_{jt} + \gamma_k \ln k_{jt} + \gamma_{kk} (\ln k_{jt})^2 + \gamma_\omega \omega_{jt} + \gamma_{\omega\omega} (\omega_{jt})^2 + \gamma_{k\omega} \omega_{jt} \ln k_{jt}, \ i > 0$
Model(1) is estimated for the sample $\mathbf{U}_a$
Model(2) is estimated for the sample $\mathbf{U}_{na}$
Model(3) is estimated for the sample $\mathbf{U}$

Table 15: Demand, cost and productivity evolution estimates

Parameter	Estim(S.E.)
$\frac{\eta+1}{\eta}$	.63(.019)*
β_k	-.29(.019)*
ρ_0	.008(.003)*
ρ_1	.99(.01)*
ρ_2	.0007(.008)
$SE(\xi)$	.09
Sample size	2828

The above estimates imply that the exponent of capital in the profit function is $(\eta+1)\beta_k = .48$, the demand elasticity is $\eta = -2.68$, and the implied markup over marginal cost is $-\frac{1}{\eta+1} = .6$.

Table 16: Reduced-form profit function

Variable	Coefficient(St. Error)
Log Capital ($\ln k$)	.84(.02)*
Sample size	2867

OLS regression with log profit ($\ln \pi$) as a dependent variable and log capital ($\ln k$) and a full set of time dummies as independent variables.
The above estimate implies that the exponent of capital in the profit function is .84, which is much higher than the estimate of .48 derived from the structural where productivity is controlled for.

Table 17: ω of Subsidized vs. unsubsidized firms

Subsidy Status Breakdown		Median ω	Mean ω	Obs.
Never Subsidized firms		0.23	0.33	1,897
Subsidized	Before ($t < t_0$)	0.32	0.35	614
	After ($t \geq t_0$)	0.41	.46	317

Table 18: Logit Estimates of the Probability to get Subsidized

Variable	Coeff.(S.E)
β_{p_0}	-3.9(.16)*
β_ω	.64(.25)*
Sample size	2828

The probability to be subsidized in location l conditional on productivity ω is given by:
$P_S(l,\omega) = exp(\beta_{p_0} + \beta_\omega \omega)/(1 + exp(\beta_{p_0} + \beta_\omega \omega))$ and is estimated with maximum likelihood.

Table 19: Empirical distribution of $G_0(p_k)$, $p_k < 1$

p'_k	.5	.55	.6	.65	.70	.8
$\hat{G}_0(p'_k)$	.17	.28	.18	.13	.17	.07
Firm-Year Obs.	2828					

By definition $\hat{G}(p'_k) = \frac{\hat{F}_{p_k}(p'_k|1,\omega)}{\sum_{p'_k<1} \hat{F}_{p_k}(p'_k|1,\omega)}$ $\forall \omega$, $\forall p'_k < 1$.

Table 20: Dynamic parameters estimates

Parameter	c_2	λ_k	σ^2_ϵ
Estimate	2e-7*	.34	.04
Firm-Year Obs.	2824		
Log-likelihood	419†		

The calibrated parameters of the model are the depreciation (δ) set at .05 and the discount factor (β) set at .95.
† For such low variance it is not unusual for the log-likelihood to be positive.

Table 21: Capital stock

Dataset	Asset Class	Deflator	δ *
ASM	Land	CPI	.000
	Buildings	GDP Non-Residential Constr.	.033
	Transport	GDP Transp. Equipment	.168
	Machinery	GDP Machinery	.109
	Remain. Equip. & Intangib.	GDP Machinery	.215†
ICAP	Land	CPI	.000
	Machinery	GDP Machinery	.109
	Intangibles	GDP for all Investment	.315
	Trans. & Build. & Equip.	GDP for all Investment	.052††

* Depreciation rates are industry specific.
† Mean between the communications equipment deflator and the intangibles deflator.
†† Data-specific. Generated by a weighted average of buildings and transport deflator.

Figure 3: Total cash subsidies granted and median subsidy rate by year

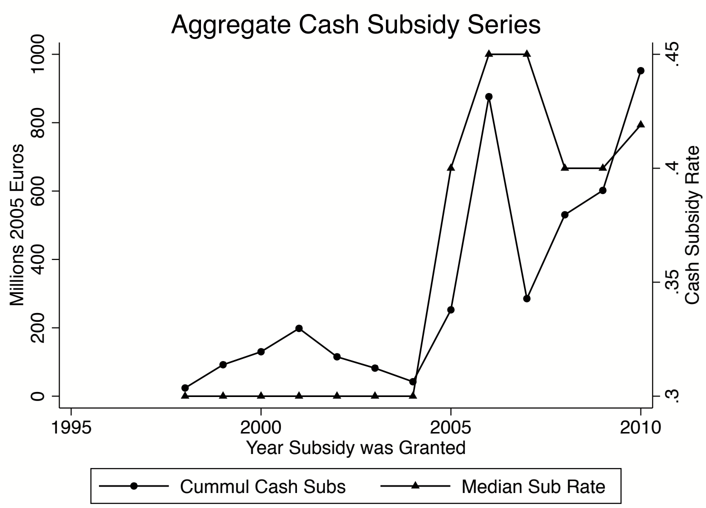
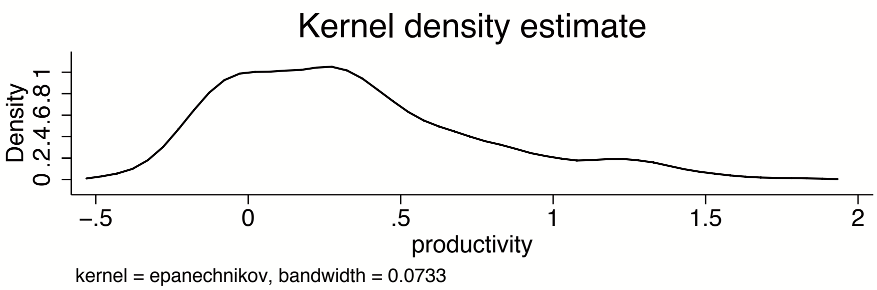
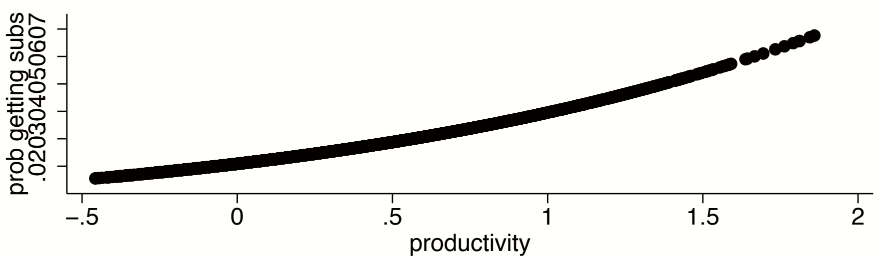


Figure 4: Unconditional distribution of productivity and probability of being subsidized





Logit predicted probabilities of receiving a subsidy conditional on productivity

Figure 5: Contribution to aggregate investment by productivity decile

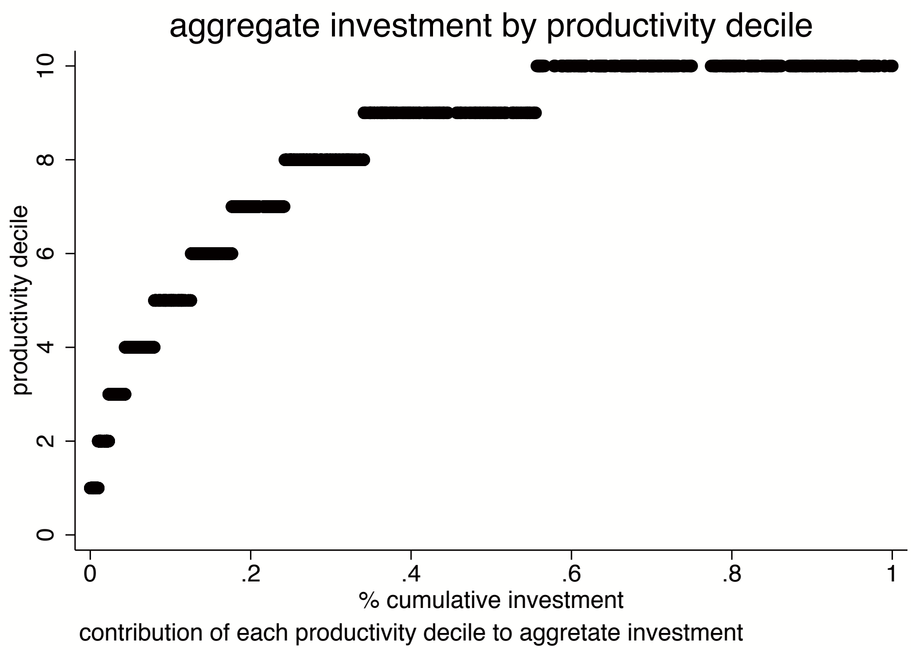


Figure 6: Optimal Investment policy. Subsidy rate 50%

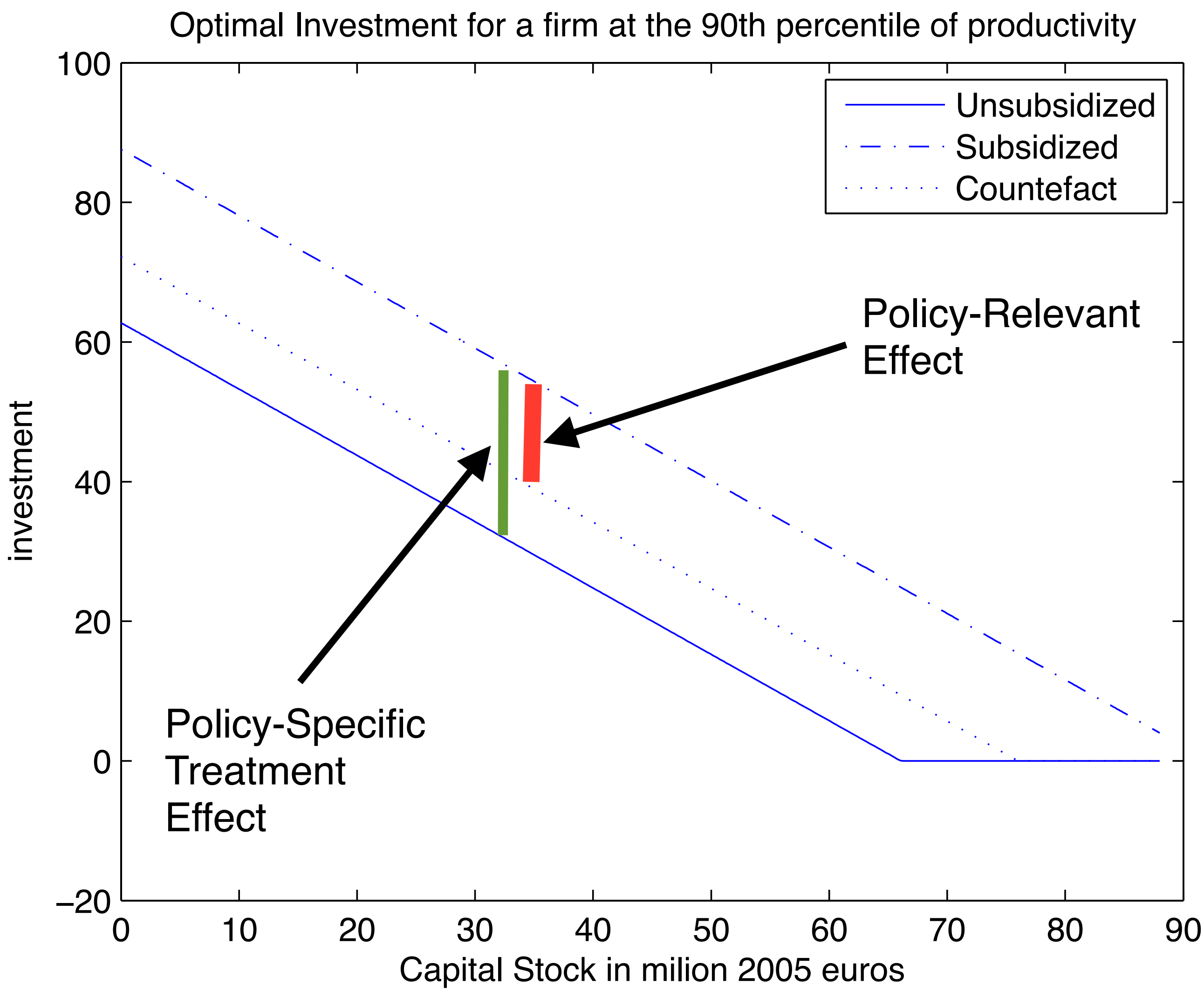
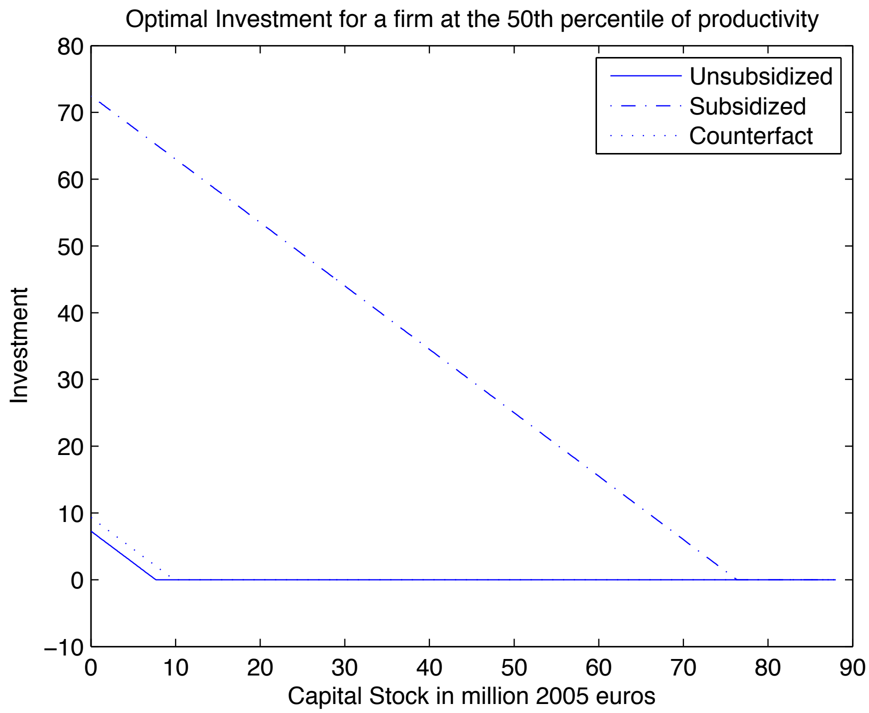


Figure 7: Optimal Investment policy. Subsidy rate 50%



D Numerical Implementation

In this section I describe the specifics of the numerical implementation of the maximum likelihood estimation in a computer. In models where the value function $V(s; c_2, \lambda_k)$, implicitly defined by the Bellman equation (9) has a closed form solution, the maximum likelihood problem defined in (33) can be solved as an unconstrained optimization problem in $c_2, \lambda_k, \sigma_\epsilon$ by most non-linear optimization software. Unfortunately, in the model of this paper, V has no closed form solution and thus, needs to be approximated by solving the functional equation (9) in some finite dimensional space. This method describes a novel numerical algorithm to solve the maximum likelihood problem efficiently.

D.1 Value Function Approximation

I approximate the value function $V(s)$ by Chebyshev polynomials of the first kind. Let $T_o(z)$ be the oth order Chebyshev polynomial $T_o : [-1, 1] \to \mathbf{R}$. I approximate the value function of dimension d by products of univariate functions T_o. For each dimension j of V, I need to choose the maximum polynomial order m_j. Let $\mathcal{T}_a(z)$ denote the product $T_{a_1}(z_1) \cdots T_{a_d}(z_d)$. To characterize the dimension the approximation space of V, I need to choose the maximum total polynomial degree $\bar{m}$. The total degree of polynomial $\mathcal{T}_a$ is defined as $|a| = a_1 + a_2 + \cdots + a_d$. The approximation space is then spanned by the basis functions $\mathfrak{T} = \{\mathcal{T}_a | 0 \leq a_j \leq m_j,\ |a| \leq \bar{m},\ 1 \leq j \leq d\}$ and its dimension is given by the cardinality of $\mathfrak{T}$. When $\bar{m} = \max\{|a| \mid 0 \leq a_j \leq m_j,\ 1 \leq j \leq d\}$, then the approximation space is a tensor product space. When $\bar{m} = m_j\ \forall\ j$ then the approximation space is a degree-$\bar{m}$ complete Chebyshev approximation space. Any function $v(z) : [-1, 1]^d \to \mathbf{R}$ in the approximation space can be expressed as a linear combination of the basis functions in $\mathfrak{T}$ that is $v(s) = \sum_{0 \leq |a| \leq \bar{m}} b_a \mathcal{T}_a(z)$. There is no general rule on choosing $\mathfrak{T}$. As a rule of thumb I start with a low cardinality $\mathfrak{T}$ and increase it until the Bellman equation holds with the desired degree of accuracy and the maximum likelihood estimates stabilize.

Since the domain of V is not the hypercube $[-1,1]^d$ a linear change of variables need to be performed to characterize the approximation space of V which I call $\mathcal{V}$. Let $s_j \in [\underline{s}_j, \bar{s}_j],\ j \leq d$ then the domain of V is $\mathcal{S} = \underset{j \leq d}{\times} [\underline{s}_j, \bar{s}_j]$. Let the function Z perform the linear change of variables $Z : \mathcal{S} \to [-1,1]^d,\ Z(s) = \left(\frac{2s_1 - \underline{s}_1 - \bar{s}_1}{\bar{s}_1 - \underline{s}_1}, \ldots, \frac{2s_d - \underline{s}_d - \bar{s}_d}{\bar{s}_d - \underline{s}_d} \right)$. Then $\mathcal{V}$ is the vector space spanned by the basis functions $\mathfrak{V} = \{\mathcal{T}_a \circ Z | 0 \leq a_j \leq m_j,\ |a| \leq \bar{m},\ 1 \leq j \leq d\}$ and any function $V : \mathcal{S} \to \mathbf{R} \in \mathcal{V}$ can be expressed as $V(s; \mathbf{b}) = \sum_{0 \leq |a| \leq \bar{m}} b_a \mathcal{T}_a(Z(s))$. Solving numerically for the value function means that I search for the Chebyshev coefficients $\mathbf{b}$ satisfying the Bellman equation (9).

I choose Chebyshev polynomials to approximate V because the firm's choice variable is continuous and optimal choices are characterized using derivatives. Hence, it is important the approximating function to be continuously differentiable and Chebyshev polynomials are by construction continuously differentiable functions[28]. Moreover, Chebyshev polynomials, have good approximation properties partly because they are orthogonal (see Judd (1998, p. 214)).

D.2 Solving the Bellman Equation

Given the approximation space $\mathcal{V}$, the Bellman equation (9) can be written as

$$\begin{aligned} V(s; \mathbf{b}) = \max_{i \geq 0} \pi(s) - ip_k - C(k, k') + \beta E_F V(s'; \mathbf{b}) \\ k' = i + (1 - \delta)k \end{aligned} \tag{37}$$

where the explicit dependence of V on $\mathbf{b}$ serves to remind that $\mathbf{b}$ defines the value function in $\mathcal{V}$. Equation (37) should hold for every $s \in \mathcal{S}$ but in order to solve the Bellman equation in a computer it is necessary to reduce (37) to a finite number of equations by requiring it

[28] Unlike most finite element methods often used for approximating multidimensional functions

to hold at a grid of points $\mathbb{S} \in \mathcal{S}$. Approximation theory tells us that there are efficient ways to choose $\mathbb{S}$ and one such way that I use is grids formed by the tensor product of expanded Chebyshev approximation nodes (see Judd (1998, p. 222)). Last, I need to approximate the expectation with respect to the continuous distributions in (37) by a finite sum which I achieve by using Gauss-Hermite quadrature (see Judd (1998, p. 261)). Consequently, the problem I solve in the computer is

$$\begin{aligned} V(s;\mathbf{b}) = \max_{i\geq 0} \pi(s) - ip_k - C(k,k') + \beta \sum_j w_j V(s_j^{-k}, k'; \mathbf{b}) \\ k' = i + (1-\delta)k, \quad s \in \mathbb{S} \end{aligned} \tag{38}$$

where w_j are weights determined by a combination of quadrature weights and the discrete probability distribution of the price of capital and s_j^{-k} are integration nodes of the stochastic state variables determined by the finite support of the distribution of the price of capital and the quadrature abscissae.

To solve equation (38) I need to have a method to solve the maximization problem in the right hand side for each of the N_g grid points in $\mathbb{S}$ given a value function $V(\cdot, \mathbf{b})$. There are two ways to solve these optimization problems. The first is to solve them separately for each grid point by calling a non-linear optimizer N_g times. The other is to replace each optimization problem with the corresponding Karush-Kuhn-Tucker (KKT) conditions[29] and solve the resulting equalities and inequalities with a non-linear equation solver. The advantage of the second method is that problem (38) becomes a set of equalities and inequalities and the maximum likelihood problem in (33) can be transformed to a non-linear constrained optimization problem which can be solved efficiently with most non-linear optimization software. The reason why this is an advantage will be discussed in the next section.

[29]The KKT conditions are sufficient for optimality since the value function is concave.

There are two ways to solve for the vector of coefficients **b** defining the value function. The first is to exploit the contraction property of the Bellman operator and use value function iteration. The other is to solve the N_g equations in (38) by a non-linear equation solver using Newton's method. The advantage of the first method is it's robustness: value function iteration almost always converges irrespective of starting values. Its disadvantage is its speed which is often too slow to be used in an estimation algorithm where the value function has to be repeatedly calculated for different parameter values. The advantage of the second method is it's speed because it utilizes derivative information but its disadvantage is its lack of robustness due to the sensitivity of Newton's method to staring values. My estimation algorithm described in the next section uses value function iteration to provide good starting values for Newton's method which subsequently solves for the value function.

Even though in theory the value function is concave, in practice the equioscillation property of Chebyshev approximation often results in a non-concave value function satisfying (38). A non-concave value function poses serious problems for my algorithm since the sufficiency of the KKT conditions crucially depends on the concavity of the value function. To tackle this problem I use shape preserving value function approximation which imposes concavity restrictions on the value function (see Cai and Judd (2010, 2013, 2014); Cai, Judd, Lontzek, Michelangeli and Su (2013)).

D.3 Maximum Likelihood Estimation Algorithm

The maximum likelihood problem in (33) is a constrained optimization problem with the Likelihood function as the objective function and the Bellman equation as the constraint. After choosing an appropriate finite dimensional space to solve the Bellman equation in, described in sections D.1 and D.2 above, the numerical version of the maximum likelihood problem becomes

$$
\begin{aligned}
\max_{\theta} \quad & \mathcal{L}(\{\tilde{s}_{it}, k'_{it}, \tau_{it}, p^h_{k_{it}}\}_{it}; \theta) \\
\text{subject to:} \quad & V(s; \mathbf{b}, \theta) = \max_{i \geq 0} \pi(s) - ip_k - C(k, k') + \beta \sum_j w_j V(s_j^{-k}, k'; \mathbf{b}, \theta) \\
& k' = i + (1 - \delta)k, \quad s \in \mathbb{S}, \quad \theta = (c_2, \sigma_\epsilon, \lambda)
\end{aligned}
\tag{39}
$$

Notice that I add θ as an argument in the value function to stress its dependence on the adjustment cost and persistence parameters c_2, λ. There are mainly two approaches to solving estimation problems like (39). One approach is to use a non-linear solver to maximize the likelihood $\mathcal{L}$ with respect to parameters θ and every time the solver evaluates the Likelihood at a point θ the Bellman equation has to solved as in Rust (1987). Another approach is to formulate (39) as a constrained optimization problem directly for the solver where the N_g equations in (38) are the constraints of the optimization problem and $\mathcal{L}$ is the objective function as in Su and Judd (2012). The advantage of the first method is its robustness since any maximum likelihood problem subject to a Bellman equation can be solved in this way but its disadvantage is its speed since the Bellman equation has to be solved hundreds if not thousands of times. The advantage of the second method is its ability to exploit the power of modern non-linear optimization software which handle the sparsity pattern of the problem efficiently and exploit derivative information efficiently. Its disadvantage is that not all maximum likelihood problems can be easily supplied to a solver in the form of a constrained optimization problem. More specifically, as Su and Judd (2012) showed, maximum likelihood estimation of discrete choice problems can be recast as a constrained optimization problem for non-linear solvers. This is mainly because, under certain conditions (see Rust (1987, 1988)), the Bellman Equation in (39) that guarantees that the agent behaves optimally, can be substituted by another functional equation that lacks the max operator. Writing the Bellman equation as a set of equality and inequality constraints is also possible for continuous choice problems when the KKT conditions of the problem are sufficient for optimality. This is true in the model developed in this paper since the adjustment cost is convex and the value function is concave and thus, the maximum likelihood problem in (39) can be written as

$$\max_{\theta,\mathbf{b},\{k'_n,u_n\}_{n=1}^{N_g}} \mathcal{L}(\{\tilde{s}_{it}, k'_{it}, \tau_{it}, p^h_{k_{it}}\}_{it};\theta)$$

subject to:

$$V(s_n;\mathbf{b},\theta) = \pi(s_n) - (k'_n - (1-\delta)k_n)p_k - C(k_n,k'_n) + \beta\sum_j w_j V(s^{-k}_{nj}, k'_n;\mathbf{b},\theta) \tag{40a}$$

$$\begin{aligned} &-p_k - \tfrac{\partial C(k_n,k'_n)}{\partial k'} + \beta\sum_j w_j \frac{\partial V(s^{-k}_{nj}, k'_n;\mathbf{b},\theta)}{\partial k'} + u_n = 0 \\ &u_n \geq 0,\ k'_n \geq (1-\delta)k_n,\ u_n(k'_n \geq (1-\delta)k) = 0 \end{aligned} \tag{KKT}$$

$$\frac{\partial^2 V(s_q;\mathbf{b},\theta)}{\partial k'^2} \leq 0,\ \ s_q \in \mathbb{Q} \tag{40b}$$

$$s_n \in \mathbb{S}, \quad \theta = (c_2, \sigma_\epsilon, \lambda)$$

Notice that the max operator in the Bellman equation in (39) was replaced by the (KKT) conditions and that a set of inequality constraints (40b) was added to guarantee that the value function is actually concave and therefore, the (KKT) conditions are sufficient for optimality. The algorithm usually works best when the grid of points for the concavity restrictions $\mathbb{Q}$ is more numerous than $\mathbb{S}$ (see Cai, Judd, Lontzek, Michelangeli and Su (2013)). The optimization problem (40) is referred to as a mathematical program with complementarity constraints (MPCC) or a mathematical program with equilibrium constraints (MPEC) and what differentiates it from standard constrained optimization problems is the presence of the complementary slackness conditions in (KKT) which make the problem hard to solve. This is because standard constraint qualification conditions, usually assumed and required by non-linear solvers to be effective, fail to hold in complementarity problems (see Ralph (2008)). Fortunately, there exist optimization software, such as KNITRO 9 I use, which contain algorithms specifically designed to solve MPCCs and make the solution of problem (40) possible.

A drawback of the MPCC formulation of the maximum likelihood problem and its solution with Newton's method is that Newton's algorithm can be sensitive to staring values if they are too far away from the solution. To ameliorate this problem, I use value function iteration to solve for $\mathbf{b}_0, \{k'_{n0}, u_{n0}\}_{n=1}^{N_g}$ at the starting value θ_0 and provide $\mathbf{b}_0, \{k'_{n0}, u_{n0}\}_{n=1}^{N_g}, \theta_0$ to the solver as a starting value. The algorithm performs well and solves a problem with 4 state variables usually in less than a day.