Solving Models with Jump Discontinuities in Policy Functions

Gortz, Christoph; Mirza, Afrasiab

DOI:
10.1111/obes.12203

Solving Models with Jump Discontinuities in Policy Functions.*

Christoph Görtz[†] and Afrasiab Mirza[‡]

This version: April 2017.

Abstract

We compare global methods for solving models with jump discontinuities in policy functions. We find that differences between Value Function Iteration (VFI) and other methods are economically significant and Euler equation errors fail to be a sufficient measure of accuracy in such models. VFI fails to accurately identify both the location and size of jump discontinuities, while the Endogenous Grid Method (EGM) and the Finite Element Method (FEM) are much better at approximating this class of models. We further show that combining value function iteration with a local interpolation step (VFI-INT) is sufficient to obtain accurate approximations. The combination of computational speed, relatively easy implementation and adaptability make VFI-INT especially suitable for approximating models with jump discontinuities in policy functions: while EGM is the fastest method it relatively complex to implement; implementation of VFI-INT is relatively straightforward and it is much faster than FEM.

Keywords: Dynamic Equilibrium Economies, Non-Convex Capital Adjustment Costs, Computational Methods, Nonlinear Solution Methods, Euler equation errors.

JEL Classification: C63, C68, E37.

11356 words.

---

*We thank Francesco Zanetti and two anonymous referees for their constructive comments which significantly improved the paper. We further thank Christian Bayer, Andrew Clausen, Wouter den Haan, Giulio Fella, John Fender, Jesus Fernandez-Villaverde, Tom Holden, Julia Iori, Kaushik Mitra, Christopher Otrok, Morten Ravn, Pontus Rendahl, Peter Sinclair, Carlo Strub, Konstantinos Theodoridis, Håkon Tretvoll, John Tsoukalas, Fabio Verona and participants at the Society of Computational Economics 2013 Conference and the Royal Economic Society Annual Meeting 2014 for useful comments and suggestions. All remaining errors are our own. Görtz acknowledges support from a British Academy Small Grant.

[†]University of Birmingham, Department of Economics, J.G. Smith Building, Edgbaston, Birmingham, B15 2TT. Email: c.g.gortz@bham.ac.uk.

[‡]University of Birmingham, Department of Economics, J.G. Smith Building, Edgbaston, Birmingham, B15 2TT. Email: a.mirza@bham.ac.uk.

1 Introduction

We examine differences in the answers produced by global approximation methods for solving dynamic economies where agents face non-concave problems (i.e. non-convex choice sets). Non-concave problems can result from the inclusion of fixed adjustment costs that are empirically relevant in many circumstances.[1] In such problems, agents make discrete decisions by comparing the option values associated with different adjustments. Fixed adjustment costs generate kink(s) in the value function at the intersection of these option values and imply jump discontinuities in the policy function. While differences across approximation methods have been extensively studied for dynamic economies where policy functions are continuous (e.g. McGrattan (1996), Santos (2000), Aruoba et al. (2006), Santos and Peralta-Alva (2012)), the literature provides little guidance about the adequacy and accuracy of computational methods when policy functions exhibit jump discontinuities. The goal of this paper is to fill this gap.

We document that the exact intersection of the option values — and thereby the location of a jump discontinuity in the policy function — is difficult to determine using discretized Value Function Iteration (VFI). The use of a finite grid on state and control variables limits VFI to approximating the option values as step functions. This results in multiple intersections of these values and leads to an imprecise determination of the jump discontinuity. Such imprecision may lead to economically significant approximation errors. Sufficient mitigation of this problem requires very fine grids that are infeasible in many applications due to the curse of dimensionality.

[1] The relevance of fixed adjustment cost is highlighted for example in studies of investment (e.g. Caballero et al. (1995), Doms and Dunne (1998), Power (1998), Cooper et al. (1999), Nilsen and Schiantarelli (2003) and Cooper and Haltiwanger (2006), Whited (2006), Bayer (2006), Khan and Thomas (2008), Bloom (2009), Wang and Wen (2012)), consumer-durables choice (e.g. Jose Luengo-Prado (2006), Bajari et al. (2013)), portfolio choice models with transaction costs and asset prices (e.g. Vayanos (1998)), costly technology adoption (e.g. Khan and Ravikumar (2002)) and optimal dynamic capital structure choice (e.g. Hennessy and Whited (2005)).

To our knowledge the problem VFI exhibits for models with jump-discontinuities has not been documented in the literature. We explore its implications and show that a Finite Element Method (FEM) and an adaptation of the Endogenous Grid Method (EGM) can overcome this problem.[2] This is essentially because both methods approximate the option values over the entire state space using piece-wise linear functions — effectively approximating these values using an infinite set of points — leading to a single intersection of option values and therefore a unique determination of the jump discontinuity in the policy function. We also show that extending VFI to allow the option values to be approximated *locally* around each grid point using piece-wise linear functions (VFI-INT) is sufficient to obtain a unique intersection and precise solutions.

We illustrate differences across methods for non-concave problems using a partial equilibrium model of a plant where investment is subject to both variable and fixed capital adjustment costs. This model is well established in the literature and is based on Cooper and Haltiwanger (2006).[3] Their paper provides widely used parameter estimates and statistics on the importance of capital adjustment costs and relies on VFI as an approximation method. In this model, in the presence of fixed costs the plant determines its investment strategy each period by comparing the option value of remaining inactive (not investing) with the option value of becoming active (investing). The optimal investment strategy follows an (S, s) adjustment process whereby the plant does not make *any* investment until capital depreciates below a threshold level at which point the plant makes a *substantial* investment to re-build its capital stock (investment spike). The threshold is determined by the intersection of the plant's option values. To correctly capture the dynamics of investment it is crucial to determine this threshold accurately. We show that EGM, FEM and VFI-INT yield a unique threshold, while in contrast, even for fine grids VFI yields multiple thresholds

[2] Given that we consider non-concave problems, we focus on piece-wise linear approximations and do not implement methods that involve higher order polynomial approximations.

[3] We illustrate these differences also in a general equilibrium setting in online appendix A.8.

located across a wide range of capital values.

We also highlight that relying on the use of Euler equation errors alone is insufficient to assess the accuracy of approximation for models with jump discontinuities in policy functions. We show that standard measures in the literature such as average or maximum Euler equation errors fail to indicate how well the threshold is approximated. To assess how well the four different methods approximate the threshold and the size of the jump discontinuity we conduct a simulation exercise and focus on two key statistics: the size of investment spikes and firm's average capital stock. These are very sensitive to the location and size of the discontinuity and are also frequently reported as key statistics in models with (S, s) behavior.[4]

We find that VFI generates statistics that are noticeably different from the "true" investment spike size and average capital stock.[5] This is in stark contrast to the performance of EGM, FEM and VFI-INT which deliver statistics very close to the true ones. Crucially, the differences between these methods and VFI are economically significant. For example the maximum percentage deviations across shocks are much higher using VFI: for a particular comparable grid, VFI implies that a firm's investment spike size (mean capital stock) deviates up to 16% (8%) from the true size. In contrast VFI-INT only implies a maximum deviation of 4% (2%). The performance of VFI is also inferior to the other three methods if a welfare based measure is used.

The limited informativeness of Euler equation errors in determining the accuracy of solutions for models with jump discontinuities in policy functions, implies that the conventional speed comparisons across methods – that use Euler equation errors to benchmark accuracy – can be misleading. We provide a first indication on the relative speed of VFI, FEM, EGM

[4] Statistics resulting from simulations have been used in the literature as an alternative to Euler equation errors to assess accuracy of approximations across methods, see for example Heer and Maußner (2008).

[5] We define the "true" statistics as the mean of those generated by EGM and FEM for very fine grids (defined in detail in Section 5) as the model does not have an analytical solution.

and VFI-INT for models with jump discontinuities in policy functions without relying on Euler equation errors alone. We compare speed by benchmarking accuracy based on the statistics introduced above. Surprisingly, we find – in contrast to the literature benchmarking on Euler equation errors – VFI is much slower than FEM. For the other methods, FEM is the slowest method followed by VFI-INT while EGM is the fastest.

However, EGM is also by far the most complex method to implement as it requires a number of adaptations to be applicable to our model. The original EGM algorithm introduced by Carroll (2006) is limited to smooth models with at most one control and one endogenous state variable. The literature has proposed numerous extensions to accommodate more complex models as the applicability of EGM is context dependent. A number of extensions have been developed recently to allow EGM to be applied to more complex models.[6] The implementation of EGM for non-smooth and non-concave problems such as ours adds a significant layer of complexity. Fella (2014) shows how to extend EGM to such settings using a consumption model that involves fixed adjustment costs for durable goods. We adapt the algorithm to our model of a plant with fixed capital adjustment costs. This problem involves an endogenous continuous choice variable that is subject to fixed adjustment costs, unlike Fella (2014), where this choice is discrete.

FEM and VFI-INT are of similar implementation complexity and are far less complex to implement for non-concave problems than EGM.[7] Both are general purpose methods that require only minimal changes to handle more complex models. However, a key drawback of FEM is that it is far more expensive in terms of computation time than VFI-INT. The

[6] These extensions often combine EGM with VFI. Barillas and Fernandez-Villaverde (2007) show how to introduce additional control variables; Hintermaier and Koeniger (2010) demonstrate how to introduce additional endogenous state variables in a durable goods model and Ludwig and Schön (2013) show how to accommodate additional endogenous state variables in a human capital model.

[7] Our FEM code approximates the value function using piece-wise linear functions with weights updated via iteration on the Bellman operator rather than minimization of the Galerkin residual as in McGrattan (1996) and Aruoba et al. (2006). The latter approach has been shown to work well for smooth problems while for our context with jump discontinuities in the policy function we find that this approach is problematic as results are highly dependent on the algorithm's start values.

combination of computational speed and easy implementation and adaptation make VFI-INT ideal for approximating models with jump discontinuities in the policy functions.

The rest of the paper is organized as follows. The next section presents the model we use to illustrate our results. We then provide descriptions of the methods we use to solve the model. Section 4 discusses the model parameterization. Section 5 analyses the differences in the solutions across methods. The final section concludes.

2 The Model

We consider a general class of models where in every period the agent makes both a continuous and discrete choice (c', d') based on the state (c, d) consisting of previous period's choices. The set of possible states is denoted by Ω. The agent's choice set is constrained as follows:

$$(c, c', d, d') \in \Gamma,$$

where Γ is $\mathbb{R}_+^2 \times \mathbb{D}^2$ where $\mathbb{D} \subset \mathbb{R}_+$ is a finite set of discrete choices.[8] Importantly, this specification of the constraints includes the case where c or d are subject to non-convex adjustment costs. The agent solves the following dynamic programming problem:

$$V(c, d, A) = \sup_{(c',d') \in \Gamma(c,\cdot;d,\cdot;A)} u(c, c'; d, d'; A) + \beta \sum_{A' \in \mathcal{A}} \pi(A'|A) V(c', d', A')$$

where $\mathcal{A}$ is the set of all possible shock realizations $A \in \mathcal{A}$, π is the corresponding transition matrix, the domain of V is $\Omega \times \mathcal{A}$, the per-period utility function of the agent is u, and the discount-factor is β. We assume that $u(\cdot, c'; d, d'; A)$ and $u(c, ; d, d'; A)$ are differentiable on $\text{int}(\Gamma(\cdot, c'; d, d'; A))$ and $\text{int}(\Gamma(c, \cdot; d, d'; A))$, respectively; and $u(\cdot)$ is an increasing and strictly concave function. Importantly, the value function V is non-concave in the presence

[8] We define particular subsets of Γ as follows: $\Gamma(c, \cdot; d, \cdot) = \{(c', d') : (c, c', d, d') \in \Gamma\}$, $\Gamma(c, \cdot; d, d') = \{c' : (c, c', d, d') \in \Gamma\}$, $\Gamma(\cdot, c'; d, d') = \{c : (c, c', d, d') \in \Gamma\}$.

of non-convex adjustment costs to c or d. As a result, the agent compares the *option values* associated with choices of c' and d'.[9] A kink in the value function arises at the point of indifference between these options and implies a jump discontinuity in the policy functions (see e.g. Clausen and Strub (2012)).

The general framework described above nests a number of important models with jump discontinuities used in the literature. This includes models with costly technology adoption (e.g. Khan and Ravikumar (2002)), durable consumption goods (e.g. Bajari et al. (2013)) and firm-level investment (e.g. Cooper and Haltiwanger (2006), Wang and Wen (2012)). We illustrate the applicability of different approximation methods using a model that captures key elements of models in the firm-investment literature. Specifically, we employ a partial equilibrium model of a plant in which capital adjustment is subject to both fixed and variable adjustment costs. It is based on a specification in Cooper and Haltiwanger (2006) which we describe in detail below.[10]

The plant produces output Y_t via the production function

$$Y_t = A_t K_t^{\alpha}, \qquad 0 < \alpha < 1, \tag{1}$$

where K_t denotes capital and productivity A_t evolves according to the AR(1) process

$$\log A_{t+1} = \rho \log A_t + \varepsilon_t, \qquad 0 < \rho < 1, \tag{2}$$

where $\varepsilon_t \sim N(0, \sigma_\varepsilon)$. The plant's capital stock evolves according to the law of motion

$$K_{t+1} = (1-\delta) K_t + I_t, \qquad 0 < \delta < 1, \tag{3}$$

[9] The dependency of the option values on the endogenous and exogenous state variables implies together with the assumption on the utility function that the option values are strictly concave.

[10] The only differences in our setup, which we applied for ease of exposition, is a simplified shock structure and irreversibility of investment.

where I_t is investment. When the plant chooses to invest, it has to pay a price p_I per investment good as well as adjustment costs $C(K_t, I_t)$. These are given by

$$C(K_t, I_t) = \frac{\gamma}{2}\left(\frac{I_t}{K_t}\right)^2 K_t + F K_t, \qquad \gamma \geq 0, F \geq 0.$$

where the first term denotes convex variable investment adjustment costs and the latter term the non-convex fixed costs. These are proportional to the capital stock to eliminate any size effects. Further, investment is completely irreversible as we assume for simplicity that capital cannot be resold on a secondary market. Formally, we impose $I_t \geq 0 \; \forall \; t$.

Note that the model includes the standard Q-theory model of investment, in which the value function is proportional to the stock of capital, as a special case.[11] The plant's problem consists of choosing a sequence of investments $\{I_t\}_{t=0}^{\infty}$ to maximize discounted life-time profits:

$$V(K, A) = \max_{\{I_t \geq 0\}_{t=0}^{\infty}} E_0 \sum_{t=0}^{\infty} \beta^t \left[A K_t^{\alpha} - p_I I_t - F \mathcal{I}_{(I_t > 0)} K_t - \frac{\gamma K_t}{2} \left(\frac{I_t}{K_t}\right)^2 \right] \tag{4}$$

subject to equations (1) and (2) and the constraint $I_t \geq 0$, given an initial level of capital, K_0, and productivity, A_0. $\mathcal{I}_{(I_t > 0)}$ is an indicator function that equals 1 if investment is positive and zero otherwise. The constraint $I_t \geq 0$ may bind in equilibrium when capital is too costly relative to the increase in future profits from additional plant capacity.

Dropping time indices, we can write the problem recursively as:

$$V(K, A) = \max\{V^a(K, A), V^i(K, A)\}, \tag{5}$$

where $V^i(K, A)$ and $V^a(K, A)$ are the values to the plant to exercising its option to either

[11] This is the case in our setup if the profits are proportional to the capital stock which is guaranteed if the plant's profit function is homogeneous of degree one ($\alpha = 1$) and the adjustment cost function is convex ($F = 0$).

remain inactive (i.e. not invest) or active (invest). We can characterize the value of the option to invest as follows:

$$V^a(K,A) = \max_{I>0} \left[AK^\alpha - p_I I - FK - \frac{\gamma K}{2} \left(\frac{I}{K} \right)^2 + \beta E_{A'|A} V(K(1-\delta) + I, A') \right], \quad (6)$$

where $K' > K(1-\delta)$. Similarly, we can characterize the value of the option to not invest as

$$V^i(K,A) = AK^\alpha + \beta E_{A'|A} V(K(1-\delta), A'), \quad (7)$$

where $K' = K(1-\delta)$ because $I = 0$. In each period, the plant computes the value of these two options and chooses its investment strategy accordingly.

In the presence of fixed costs ($F > 0$), it is optimal for the plant to follow an (S,s) adjustment strategy for investment (see Cooper et al. (1999)). In other words, investment will be zero for all periods in which the capital stock exceeds a threshold level $\hat{K}(A)$. When capital has depreciated below the threshold level the plant will make a substantial investment (i.e. undergo an investment spike) to build capital up again. Hence, there is a jump discontinuity in the policy function for investment at the threshold $\hat{K}(A)$.

The intuition behind the plant's choice is the following: for capital stock levels below $\hat{K}(A)$ the value of investing will be higher than the value of not investing: $V^a(K,A) > V^i(K,A)$. That is, the benefit from having a larger capital stock in the future exceeds the costs of investing today. For capital stocks above $\hat{K}(A)$ the opposite is true: the benefit from having an even larger capital stock tomorrow diminishes (due to decreasing returns to scale in production) and is smaller than the costs of investment. In this case $V^a(K,A) < V^i(K,A)$ and the plant will not invest.

The convexity of the adjustment costs in investment and the monotonicity of the value function $V(K,A)$ in capital entail that $V^i(K,A)$ and $V^a(K,A)$ cross exactly at one point for a given productivity, namely at $\hat{K}(A)$.[12] This implies that the value function $V(K, A)$ exhibits a kink at $\hat{K}(A)$ and is globally non-concave which is illustrated in Figure 1. As there is no closed form solution for the value function, we need to approximate the solution numerically.

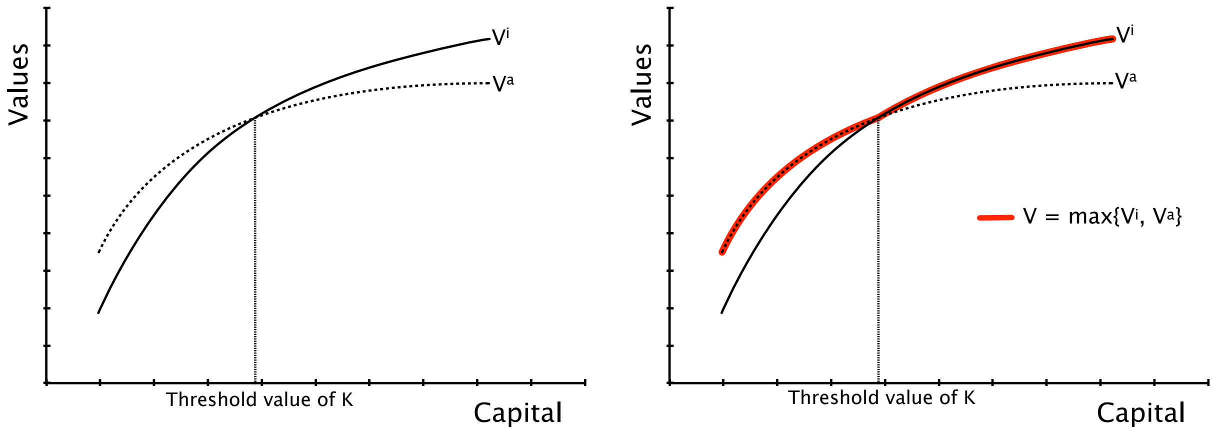

Figure 1: The left diagram shows the option values of active (V^a, dotted line) and inactive (V^i, solid line) investment for a given level of productivity. The right diagram shows the value function for the plant's problem (V, in red) that results from choosing the maximum value of the options V^a and V^i for each level of capital.

3 Solution Methods

We solve the model using four methods and provide brief descriptions of these in this section. Additional implementation details are provided in the Online Appendix. For all solution methods we approximate the AR(1) process for productivity using a discrete Markov chain as in Rouwenhorst (1995).[13]

[12] Monotonicity of the value function in capital follows from the monotonicity of the Bellman operator.

[13] We employ the method by Rouwenhorst (1995) to discretize the shock instead of the often used method by Tauchen (1986). Kopecky and Suen (2010) find that for AR(1) processes with low persistence both methods can produce highly accurate approximations, however the performance of the Tauchen method deteriorates when the serial correlation is very close to one – which is also acknoledged in the original Tauchen (1986) paper. They further find that the Rouwenhorst (1995) method outperforms Tauchen (1986) when a coarse state space is used to approximate the AR(1) process.

3.1 Value Function Iteration

The Bellman operator in our case is:

$$TV(K,A) = \max \Bigg\{ AK^{\alpha} + \beta E_{A'|A} V(A',(1-\delta)K), \\ \max_{K' \geq (1-\delta)K} AK^{\alpha} - pI - FK - \frac{\gamma K}{2}\left(\frac{I}{K}\right)^2 + \beta E_{A'|A} V(K',A') \Bigg\}, \quad (8)$$

and the laws of motion for capital and productivity are:

$$K' = (1-\delta)K + I$$

$$\log A' = \rho \log A + \varepsilon.$$

To solve the Bellman operator, we define an equally spaced grid on capital $G_K \equiv [K_1, \dots, K_n]$, and use Rouwenhorst (1995) to discretize the stochastic process for productivity A, $G_A \equiv [A_1, \dots, A_m]$. We iterate until convergence on the Bellman operator (8) to obtain an approximation of the value function over the specified grid.

This method requires the explicit computation of $V^i(K,A)$ and $V^a(K,A)$ at each grid point (root-finding step). Finally, the value function is then updated for every combination of the grid points in G_K and G_A according to

$$V(K,A) = \max\{V^i(K,A), V^a(K,A)\}.$$

When finer grids are considered, memory and computational time increase exponentially due to the need to repeatedly apply the max operator above and the use of large matrices to store the active and inactive values.

3.2 Value Function Iteration with Local Interpolation

We modify the root-finding step in VFI as follows. For every grid point (K_i, A_j), let (K^*_{i-1}, A_j), and (K^*_{i+1}, A_j) be the grid points adjacent to (resp. to the left and right of) the optimal choice of K' found by VFI for a given productivity A_j.[14] To increase the accuracy of the approximation, we generate new capital grid points on the intervals $[(K^*_{i-1}, A_j), (K^*_i, A_j)]$, and $[(K^*_i, A_j), (K^*_{i+1}, A_j)]$ via linear interpolation. We compute the option values (V^i_{INT} and V^a_{INT}) at these additional grid points for K', and again update the value function according to $V(K_i, A_j) = \max\{V^i_{INT}, V^a_{INT}\}$ and update the policy function as the corresponding optimal value of K'. Then, we continue with the VFI algorithm and iterate on the Bellman operator (8) till convergence.

3.3 Finite Element Method

The main idea behind FEM is to approximate a function of interest using a number of much simpler basis functions. Each of these basis functions are typically non-zero only on a small part of the state space, or equivalently on a small number of elements. This sparsity allows a large number of elements to be handled and the algorithm is well suited for parallel computing.

Our FEM algorithm approximates the value function using a piece-wise linear approximation. We partition the state space into rectangles of the form $[K_i, K_{i+1}] \times [A_j, A_{j+1}]$. We then approximate the value function over the state space using a piece-wise linear function over the grid points of the partition. Given an initial guess for the value function $V^0(K, A)$ at each grid point in the state space, we approximate the value function as $\hat{V}(K, A) = \sum_{ij} \hat{V}_{ij}(K, A)$

[14] Where we denote K_i (A_j) as the i^{th} (j^{th}) grid point for capital (the shock).

where

$$\hat{V}_{ij}(K,A) = \begin{cases} V^0_{ij}(K,A) + \frac{V^0_{i+1j}-V^0_{ij}}{K_{i+1}-K_i}(K-K_i) \text{ if } K \in [K_i, K_{i+1}] \\ 0 \text{ otherwise} \end{cases} \tag{9}$$

so that we effectively use a piece-wise linear approximation for each value of productivity. We then apply the Bellman operator (8) using $\hat{V}(K', A')$ as our guess for tomorrow's value function and update our initial guess $V^0(K, A)$ on the grid points. Finally, we iterate to convergence on $\hat{V}(K, A)$.

The key difference between FEM and VFI is that with FEM tomorrow's value function can be evaluated at any point in the state space. Crucially, this implies that the optimal choice of tomorrow's capital is not restricted to be on the exogenous grid $[K_1, \ldots K_n]$. Therefore, the optimization step in the Bellman operator can be carried out using a standard constrained optimization routine that enforces the irreversible investment constraint $I \geq 0$. Hence, FEM permits an additional degree of freedom above VFI but it comes at a cost as we are forced to employ the computational expensive constrained optimization routine repeatedly. Note that the policy function generated by this procedure is also a piece-wise linear function akin to (9). Our algorithm for FEM is no more difficult to implement than VFI given that we do not rely on Galerkin weighting and use standard methods for implementing root-finding such as Golden Section Search.[15]

3.4 Endogenous Grid Method

EGM as introduced by Carroll (2006) suggests assigning an exogenous grid over the control variable K' rather than the state variable K. Then, using the following first-order condition allows us to determine an endogenous grid over K, given the exogenous grid K'

[15] For our case of jump discontinuities in the policy function we find that Galerkin weighting is not suitable as it leads to results that are highly dependent on start values for the algorithm.

and the derivative of the value function with respect to K', $V_{K'}(K', A')$,[16]

$$p_I + \gamma \frac{K' - (1-\delta)K}{K} = \beta E_{A'|A} V_{K'}(K', A'). \tag{10}$$

Interpolating the solution on the endogenous grid, to evaluate it on the exogenous grid, we can obtain a set of optimal control and state pairs that can then be used to approximate the value function.

Crucially, this procedure requires a unique solution to the first-order condition (10) for every K' when solving for the endogenous grid over K. As shown in Figure 1, for our class of models fixed costs introduce kink(s) in the value function resulting in jump discontinuities in the (otherwise smooth and decreasing) slope of the value function, $V_{K'}(K', A')$. As a result, the first-order condition (10) does not imply a unique endogenous grid over K.[17] Therefore, EGM as introduced by Carroll (2006) is not directly applicable. Instead, we employ a modification of EGM proposed by Fella (2014) and implement the following steps for our case with fixed capital adjustment costs:

1. We begin by assigning an (exogenous) grid on K' and an initial guess for $V_{K'}(K', A')$.
2. We generate an endogenous grid for K using the first-order condition (10).[18]
3. We then split our endogenous state space into two regions: one where the value function is concave ($V_{K'}(K', A')$ is smooth) and another where the value function is not concave ($V_{K'}(K', A')$ exhibits jump discontinuities).
   (a) We apply Carroll (2006)'s algorithm in the concave region (as explained at the beginning of Section 3.4).

[16] The derivation of the first-order condition is shown in the Appendix.

[17] Clausen and Strub (2012) show that at the optimum the first-order condition holds and the envelope condition is valid.

[18] Note that this implies that the endogenous grid changes in every iteration.

(b) We apply VFI in the non-concave region to identify and retain only global optima.

4. We then proceed to interpolate over the endogenous state space and construct optimal (K, K') pairs in both the active and inactive cases.

5. We use these pairs to construct an approximation of the values to being active and inactive and thereby the overall value function.

6. We use the slope of this value function to construct the endogenous grid as in step 2. Steps 2-5 are repeated until the value function is deemed to have converged.

The computationally most demanding task of this algorithm is the interpolation step which is far less expensive than the maximization/optimization steps in VFI or FEM. Note however that the applicability of EGM is context dependent, for example it cannot necessarily accommodate additional variables. The reason is that EGM's applicability rests on finding a unique solution to the first-order condition(s) for the endogenous grids. This limits EGM to models with only one continuous choice variable. The literature shows how to accommodate additional variables for specific classes of models by often combining EGM with VFI steps (see for example Barillas and Fernandez-Villaverde (2007), Hintermaier and Koeniger (2010), Fella (2014), and Ludwig and Schön (2013)). In comparison to smooth and convex problems, implementation complexity increases even more for models with jump discontinuities in the policy function due to the need to identify and handle the non-concave region of the value function (Step 3) separately.

4 Parameterization

Our choice of the model parameters is based on estimates by Cooper and Haltiwanger (2006). Using annual plant level data of the Longitudinal Research Database, they estimate the above model with convex and non-convex capital adjustment costs and find that a combination of these fits the data well.[19] Importantly, they find evidence for substantial fixed adjustment costs of roughly 4% of the average plant-level capital stock. We approximate the stochastic process for productivity by a ten-state Markov chain using the method proposed by Rouwenhorst (1995) and calibrate the Markov chain to match the standard deviation $\sigma_\varepsilon = 0.03$ and persistence, $\rho = 0.885$. All estimates of Cooper and Haltiwanger (2006), that we use to calibrate the model, are summarized in Table 1.

We approximate the value function for all solution methods over the same state space for capital.[20] As a baseline scenario we use 700 (VFI), 97 (EGM), 95 (FEM), and 385 (VFI-INT) capital grid points, which results in state spaces which are representative for many practical applications.[21] For VFI-INT we use 35 interpolation points on each side of the optimal grid point identified by the value function iteration algorithm. This capital grid choice for the baseline scenario generates comparable average log-absolute Euler equation errors across methods (see Table 2).[22] As conventional in studies which consider the performance of different approximation methods we use an equally spaced grid for capital.

5 Results

This section first documents two specific issues when approximating models with jump discontinuities: (i) VFI fails to accurately approximate such models and (ii) Euler equation

[19]For the sake of simplicity of exposition we do not include the possibility of selling capital considered by Cooper and Haltiwanger (2006). Selling plant's capital stock at a price smaller than p_I would introduce an additional kink in the value function. The solution methods can be adjusted to accommodate the additional choice, but as our findings can be generalized to these additional kinks we assume irreversibility of capital for ease of exposition.

[20]The state space is chosen so that capital does not hit any boundaries during our simulations. Convergence is evaluated by considering the largest absolute distance between corresponding points of the value function of two consecutive iterations. If this absolute distance falls below 10^{-4} the algorithm is deemed to have converged.

[21]See for example Tsoukalas et al. (2016) and Görtz and Tsoukalas (2013).

[22]As the Euler equation is a necessary but not a sufficient condition in our setup, Euler equation error statistics are calculated for policy functions across all shocks solely in the area of the state space in which all approximation methods imply positive investment.

Table 1: Model Parameters (based on Cooper and Haltiwanger (2006))

β	0.95	discount factor
δ	0.069	capital depreciation rate
p_I	1	price to buy capital
α	0.592	returns of capital
ρ	0.885	persistence of plant specific shock
σ_ε	0.03	standard deviation of plant specific shock
γ	0.049	convex adjustment costs
F	0.039	fixed adjustment costs

errors are not a suitable measure for algorithm accuracy. Then we show in Section 5.2 via a simulation exercise that the inaccuracies resulting from (i) are economically significant. This exercise also shows that VFI-INT, EGM and FEM can address the shortcomings of VFI. In light of (ii), we provide in Section 5.3 a comparison of methods with respect to speed and implementation complexity.

5.1 Specific Problems in Models with Jump Discontinuities

Problems with VFI. As noted in Section 2, theory predicts that the policy function exhibits a jump discontinuity at the threshold separating the active (positive investment) and inactive (no investment) regions.[23] Figure 2 shows the policy functions for tomorrow's capital generated by VFI, VFI-INT, EGM and FEM for the baseline scenario. This figure highlights that VFI-INT, EGM and FEM all produce similar policy functions. However, these differ substantially from the one produced by VFI in two important aspects. First, VFI does not uniquely determine the threshold separating the active and inactive regions. Moreover, VFI does not approximate the shape of the active region accurately.

To more clearly see the problems that arise in the determination of the threshold with VFI, we show in the bottom panel of Figure 3 the values to the plant of being active and

[23] Theory also predicts additional jump discontinuities in the policy function in the active region due to the interaction between fixed and convex variable adjustment costs (see e.g. Clausen and Strub (2012)). The variable costs penalize the plant for making large adjustments while fixed costs penalize the plant for making small and frequent investments. The result is that the active region of the policy function consists of concave parts that are separated by jump discontinuities.

inactive for increasingly finer capital grids. The intersection of these values determines the capital threshold below which the plant is active and above which the plant is inactive. While theory predicts a single intersection of these functions, VFI generates multiple intersections as a result of approximating these values using step functions. The reason for these steps is that only a finite set of points can be used to approximate the values of being active and inactive because VFI limits the choices for both the values of the endogenous state and the control variable to a fixed grid.[24] VFI's inaccurate determination of the threshold can also be seen in the corresponding policy functions for tomorrow's capital which are shown in the top panel of Figure 3. From there it is evident that even a very fine grid using 3000 points does not deliver a unique intersection of the option values.[25]

In addition to this illustration, we provide a more comprehensive overview about the inaccurate determination of the threshold: we consider the percentage difference between the value for today's capital implied by the grid point $min(V^i > V^a)$ and today's capital implied by the grid point $max(V^i < V^a)$. Table 2 (column 7) shows that for the VFI baseline scenario (grid scenario 3) the mean across all shocks of this measure is 8.56, i.e. the capital stock to the right of the last intersection of the option values V^i and V^a is 8.56% higher than the capital stock to the left of the first intersection of the option values. This is equivalent to an average of 17.3 capital grid points across all shocks (column 9). The standard deviation across shocks of the percentage difference between the two capital stocks

[24] Such approximations are particularly prone to error when the slope of the underlying function is steep. In our problem, the slope of the value to being inactive is much larger than the slope of the value to being active. Hence, as shown in Figure 3, the approximation of the value of being inactive is much worse than the approximation of the value of being active.

[25] While we outlined in section 2 that we consider a class of models with strictly concave option values, we are grateful to one of our referees for pointing out that there are also models with linear rather than strictly concave option values. For example in Arellano (2008) and Arellano and Ramanarayanan (2012) the option value to default on bonds is independent of the endogenous state variable (bond holdings), which implies this option value is a horizontal function in bond holdings. The intersection of a horizontal function with the concave option value of not defaulting is always unique, even if the concave function is approximated by a non-decreasing step function (e.g. by using VFI). Approximation quality of this class of models is discussed in detail in Arellano et al. (2016).

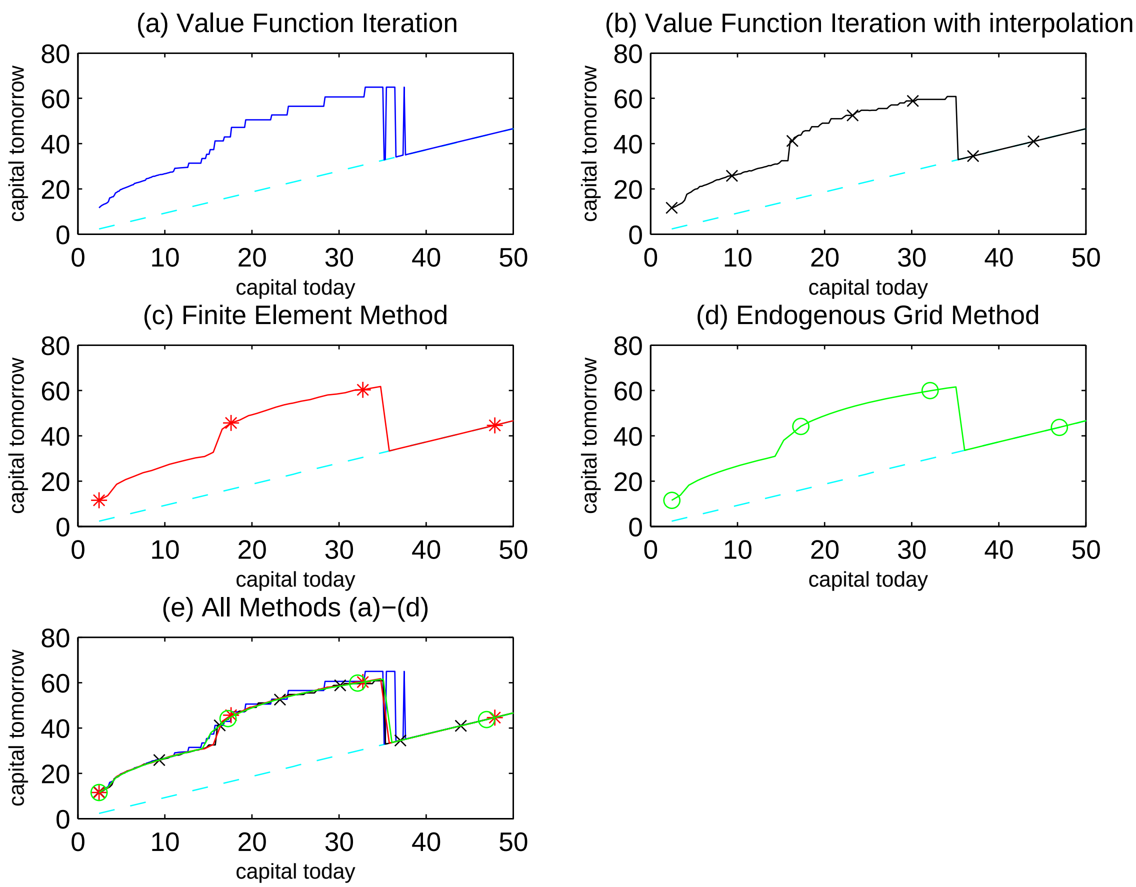

Figure 2: Policy Function for capital implied by different approximation methods. The baseline grid generates comparable average log-absolute Euler equation errors across methods. Subplots are shown for shock value 7. The blue dashed line in each subplot indicates the no investment decision $(1-\delta)K$. For better visibility we do not show part of the state space to the right of the threshold where the policy function is equal to the no-investment line.

is 5.58%. This indicates that even using 700 capital grid points (baseline scenario) the threshold is determined very imprecisely across all shocks. While the percentage difference decreases with finer grids, even for grids as fine as 1900 points (grid scenario 10) the imprecise determination of the threshold is still apparent as the mean capital stock to the right of the last intersection of the option values is 3.20% higher than the capital stock to the left of the first intersection, with a standard deviation across shocks of 1.42%. Essentially, when using VFI to approximate models with jump discontinuities in the policy function, extremely fine grids are required to determine the threshold relatively precisely. Such fine grids are typically infeasible in most applications due to the curse of dimensionality.[26]

Table 2 reports the same statistics also for VFI-INT, FEM and EGM. The grid scenarios 1-10 in this table are comparable across methods in terms of average log-absolute Euler equation errors. For the baseline scenario, VFI approximates the threshold with an imprecision that is up to 10 times larger than for other methods – 8.56% versus 0.85% (VFI-INT), 3.48% (FEM), 3.45% (EGM). The three alternative methods deliver – in line with the predictions by theory – a single intersection of V^i and V^a (the mean number of grid points across shocks is exactly unity). These methods are therefore much more suitable than VFI for approximating models with jump discontinuities in policy functions. The percentage differences in capital stocks reflect here only the distance between two adjacent grid points to the left and right of the threshold. The relatively large numbers for coarse grids highlight that in order to approximate the location of the jump discontinuity precisely, these three methods require finer grids relatively to grids that are sufficient to approximate smooth and concave models (i.e. convex choice sets).

[26]For similar reasons, VFI is also unable to correctly approximate the jump discontinuities and concave parts of the policy function in the active region. However, while the threshold is crucial for the dynamics of the model, the poor approximation of the active region is only of larger importance when the persistence of the technology shock is low.

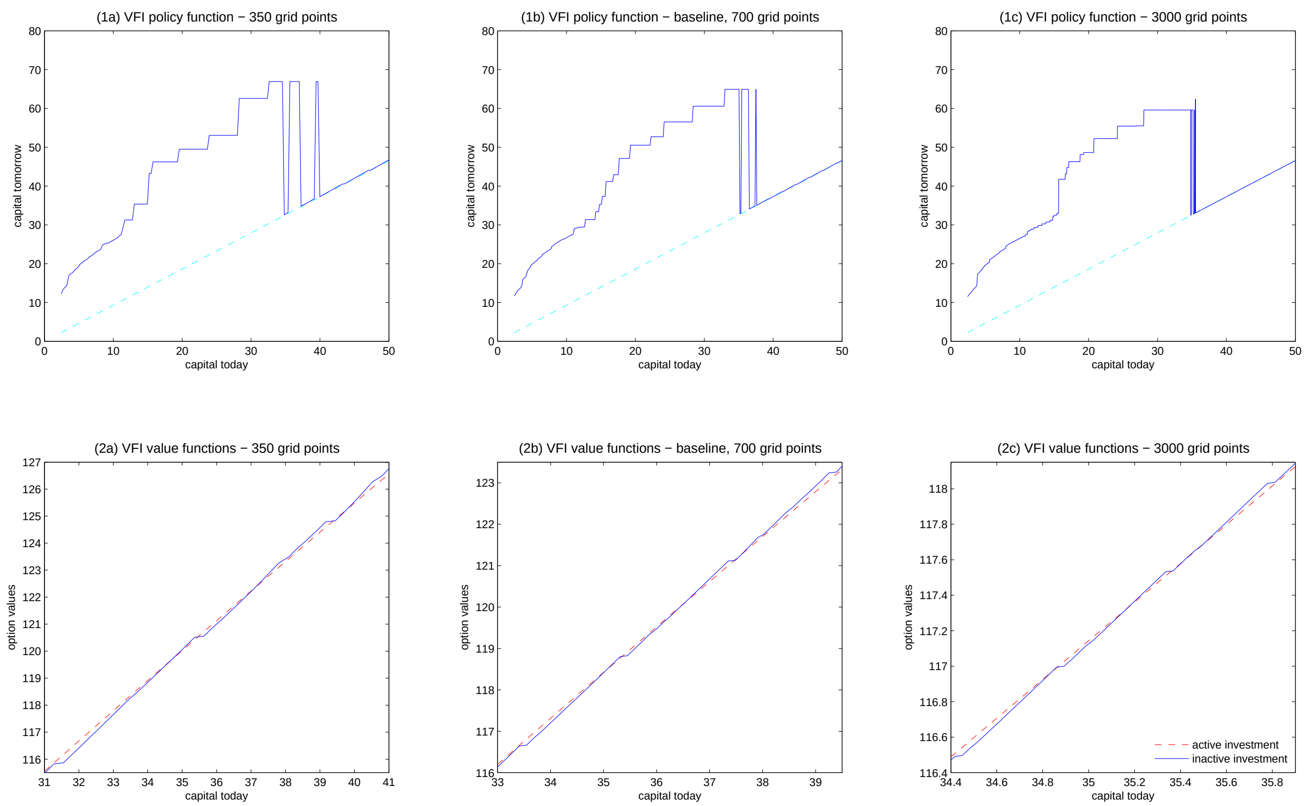

Figure 3: Approximation using Value Function Iteration. Top panel: Policy functions for capital for different grid sizes for a particular productivity level. Bottom panel: Option values to the plant of being active (red dashed) and inactive (blue solid) for different capital grids (zoomed in to show multiple intersections). Subplots are shown for shock 7.

Limited Informativeness of Euler Equation Errors. Euler equation errors are often employed as a measure of accuracy and for comparisons across methods. However, it has so far been overlooked in the literature that the information about the accuracy of approximation provided by average and maximum Euler equation errors for the class of models with jump discontinuities in the policy functions is limited. To assess accuracy of solutions for this class of models, one needs to evaluate approximation accuracy for the active region, the inactive region and the location of the threshold that separates the two regions (see Figure 1). Clausen and Strub (2012) show that Euler equations hold only in the active region for models with jump discontinuities, but they do not hold in the inactive region. Hence, Euler equation errors are a useful indicator about the quality of approximation *within* the state space that has been identified as the active region, but for this class of models, Euler equation errors cannot provide any indication about the quality of approximation (i) *within* the inactive region and (ii) of the threshold's *location* separating the two regions.

The discussion in the paragraph above highlights that the theoretical reason for limited informativeness of Euler equation errors for models with jump discontinuities is that they only hold in the active region. The points made in this discussion can also be formalized employing the model from Section 2: using equation (6) (for $V^a(K,A)$), the following first-order condition holds for the active region

$$p_I + \gamma \frac{K' - (1-\delta)K}{K} = \beta E_{A'|A} V_{K'}(K', A')$$

where $V_{K'}(\cdot)$ denotes the function's derivative with respect to K'.[27] Hence, Euler equation errors computed using this equation provide an indication about the quality of approximation of the *slope*, $V_{K'}(\cdot)$, of the value function in the active region. The Euler equation error statistic does not provide any indication about the quality of approximation of (i) the *slope*

[27]Details of the optimization problem are shown in online appendix A.1.

of the value function in the inactive region (equation (7) for $V^i(K, A)$), nor (ii) the *levels* of V^a and V^i which are key to determine the location of the threshold via the binary maximisation problem shown in equation (5).

Euler equation errors provide a useful measure for accuracy in the active region; and assessing accuracy in the inactive region does not provide a challenge as the exact solution is known.[28] However, inference about accuracy of the threshold location is less trivial. Since there is no theoretical background provided by economic models — comparable e.g. to the Euler equation for the active region — to assess approximation quality for the binary maximisation problem between V^a and V^i, one can only rely on indirect measures. Such measures are used in Section 5.2, where we suggest a simulation approach to compare key model statistics as guidance for approximation accuracy that includes a precise threshold determination.

As a result from the discussion above, Euler equation errors alone are not a sufficient measure of accuracy when policy functions exhibit jump discontinuities. While the discussion implies this is a general feature of this class of models, we also employ results from the model in Section 2 to further illustrate the matter: it is striking from Table 2 that for comparable average Euler equation errors (column 4), VFI and the other three methods deliver very different policy functions in terms of (at least) the determination of the threshold (columns 7 and 8) which is a direct implication of different approximation precisions in the binary maximisation problem.

We also compute Euler equation errors near the threshold. These are generated by simulating the model's (S, s) behavior for each shock value individually and calculating the mean of absolute log-Euler equation errors across all shocks for observations with positive investment.[29] These are reported in Table 2 (column 6) and it is evident that – as one would

[28] For the model in Section 2 recall that when the firm is inactive next period's capital stock is given by $K_{t+1} = (1 - \delta)K_t$.

[29] To clearly identify the effects of imprecise threshold determination we determine in this (and the fol-

expect from the discussion above – also the Euler equation errors at the threshold provide limited guidance on accuracy as they are very similar across grid scenarios.

We highlight even further, by considering model dynamics, that purely relying on Euler equation errors for models with jump discontinuities can be misleading. For model dynamics — which are the point of interest in most applications rather than the pure policy function — the inactive region and the threshold are much more important than the active region. We add to the discussion in Sections 2 and 5.1 on model dynamics and the importance of precise threshold determination by conducting the following exercise: we simulate our model's time series behavior for 1,000,000 periods using EGM baseline policy functions and calculate frequencies how often each grid point of the policy function is visited.[30] 90% of the time is spent in the inactive region of the policy function and 10% in the active region. This is consistent with the (S,s) behavior that implies periods of inactivity, when the capital stock just depreciates, are interrupted by periods of investment spikes. The vast majority of time spent in the active region is spent at the threshold, 83%, which stresses the importance of a precise threshold determination.[31] The other parts of the active region are only visited due to changes in the shock's value. Hence, the time spent in the inactive region depends crucially on the precise approximation of the location of the threshold and the size of the jump discontinuity which are not captured by Euler equation errors. In contrast, the time spent in the active region, about which Euler equation errors are informative in terms of approximation quality, is very limited.

lowing) simulation exercises the model's (S, s) adjustment behavior implied by a particular shock value at a time and report the mean of these exercises across shock values. We simulate the model for each shock value for 1050 periods and discard the first 50 periods to remove any impact of start values.

[30]Results are consistent across approximation methods.

[31]The threshold is defined for each policy function conditional on a shock value, as the area between the grid point closest to (1-δ) times the first grid point in the inactive region, and the last grid point of the active region.

Table 2: Statistics across different approximation methods

	grid scenario	capital grid points	Euler equation errors: average	Euler equation errors: maximum	Euler equation errors: threshold	Inaccuracy in capital threshold determination: difference in % (across shocks): capital mean	difference in % (across shocks): capital stdev.	mean number of grid points
VFI	1	350	-1.58	-0.95	-1.11	17.64	6.75	18.20
	2	600	-1.71	-0.95	-1.12	11.78	4.17	20.40
baseline	**3**	**700**	**-1.76**	**-0.95**	**-1.11**	**8.56**	**5.58**	**17.30**
	4	850	-1.78	-0.96	-1.15	7.73	3.17	19.30
	5	950	-1.80	-0.97	-1.11	6.87	2.00	20.10
	6	1050	-1.83	-0.97	-1.15	5.69	2.19	17.90
	7	1400	-1.88	-0.97	-1.11	4.04	2.00	16.80
	8	1900	-1.93	-0.98	-1.15	3.20	1.42	18.10
	9	2300	-1.96	-0.98	-1.11	2.43	1.30	16.20
	10	3000	-1.99	-0.98	-1.13	1.82	0.79	16.70
VFI_int_35	1	115	-1.59	-0.99	-1.13	2.89	0.86	1
	2	276	-1.72	-1.00	-1.13	1.18	0.34	1
baseline	**3**	**385**	**-1.76**	**-1.00**	**-1.12**	**0.85**	**0.25**	**1**
	4	400	-1.78	-1.00	-1.12	0.81	0.24	1
	5	500	-1.80	-1.00	-1.12	0.65	0.19	1
	6	750	-1.84	-1.00	-1.13	0.43	0.13	1
	7	1000	-1.88	-1.00	-1.13	0.32	0.09	1
	8	1700	-1.93	-1.00	-1.13	0.19	0.06	1
	9	2300	-1.96	-1.00	-1.12	0.14	0.04	1
	10	2500	-1.97	-1.00	-1.13	0.13	0.04	1
FEM	1	27	-1.57	-1.01	-1.12	13.35	4.46	1
	2	83	-1.72	-1.00	-1.11	4.07	1.23	1
baseline	**3**	**95**	**-1.74**	**-1.01**	**-1.13**	**3.48**	**1.01**	**1**
	4	150	-1.78	-1.00	-1.13	2.19	0.64	1
	5	175	-1.80	-1.01	-1.12	1.88	0.55	1
	6	250	-1.84	-1.00	-1.13	1.31	0.38	1
	7	330	-1.89	-1.00	-1.13	0.99	0.29	1
	8	550	-1.93	-1.00	-1.12	0.59	0.17	1
	9	800	-1.96	-1.00	-1.12	0.41	0.12	1
	10	1000	-1.98	-1.00	-1.12	0.32	0.09	1
EGM	1	34	-1.57	-1.01	-1.13	10.24	2.99	1
	2	70	-1.72	-1.01	-1.12	4.86	1.45	1
baseline	**3**	**97**	**-1.76**	**-1.01**	**-1.12**	**3.45**	**1.00**	**1**
	4	120	-1.79	-1.01	-1.11	2.77	0.81	1
	5	128	-1.80	-1.01	-1.12	2.60	0.75	1
	6	180	-1.84	-1.01	-1.12	1.83	0.53	1
	7	210	-1.88	-1.01	-1.11	1.57	0.46	1
	8	400	-1.93	-1.01	-1.12	0.82	0.24	1
	9	500	-1.96	-1.01	-1.12	0.65	0.19	1
	10	650	-1.99	-1.01	-1.11	0.50	0.14	1

VFI: Value Function Iteration, VFI-INT: VFI with local interpolation, EGM: Endogenous Grid Method, FEM: Finite Element Method. Average and maximum Euler equation errors are calculated across policy functions for all shocks in the area of the state space in which the Euler equation holds. We calculate the threshold Euler equation error as follows. For each shock value we simulate time series of 1050 periods of which the first 50 periods are discarded. For every observation we calculate the Euler equation error if it is valid. The statistics reported is the mean Euler equation error across all shocks and all simulated periods. The last three columns show statistics about the percentage difference between the value for today's capital implied by the grid point $min(V^i > V^a)$ and today's capital implied by the grid point $max(V^i < V^a)$: column 7 shows the mean percentage difference across all shocks and column 8 reports the standard deviation. Colunmn 9 shows the corresponding average number of grid points across shocks.

5.2 Economic Significance

We explore the economic relevance of the differences across methods documented above through a simulation exercise that focuses on two key statistics of the model: the size of investment spikes and the mean of capital. These two statistics are often used to calibrate models with (S, s) adjustment of capital to the data. Moreover, the mean of capital is a popular measure for firm size. These two statistics crucially depend on a precise determination of the location and size of the jump discontinuity. While the size of investment spikes and the mean of capital are useful to describe the (S, s) adjustment behavior resulting from the fixed adjustment costs, these two statistics are also specific to the class of models with this feature. Since jump discontinuities can arise in a variety of models, it is appealing to report more general statistics on welfare that can be computed for different classes of models.

The following simulation exercise focuses on the effects near the threshold as this is most important for model dynamics. We evaluate the distance between the two key statistics implied by the different approximation methods and the "true" statistics. As the model does not have an analytical solution, we solve it using FEM and EGM — the two methods that allow by construction for the highest accuracy of threshold determination — for a large number of grid points and label the average statistics produced by the resulting policy functions to be the "true" statistics.[32] We solve the model with the four approximation methods at the grid scenarios shown in Table 2. For each method and grid scenario, we simulate the (S, s) behavior of the model for a particular shock value and calculate the absolute percentage deviation of the size of investment spikes and capital mean from the corresponding true solution.[33] The average deviations of these statistics across all shock

[32]In particular we use 4000 capital grid points for FEM, and 2500 for EGM. These deliver a unique intersection of V^i and V^a. The average absolute percentage deviation between the two sets of statistics across all shocks is 0.5% for the average capital stock and 0.6% for the average investment spike size. These differences are also consistent with results we obtain using alternative grids (2400 for FEM and 1500 for EGM). Our results also continue to hold using these alternative grids.

[33]We simulate the model for each shock value for 1050 periods and discard the first 50 periods to remove any impact of start values.

values as well as the maximum deviations are shown in Table 3.

Table 3: Percentage absolute deviation across shocks from the true statistics

grid scenario	Panel 1a: investment spike size mean deviation across shocks				Panel 2a: capital stock mean deviation across shocks				Panel 3a: welfare mean deviation across shocks			
	VFI	FEM	EGM	VFI-INT	VFI	FEM	EGM	VFI-INT	VFI	FEM	EGM	VFI-INT
1	9.03	5.27	3.30	5.35	4.76	2.49	2.68	1.78	2.61	1.62	2.99	1.12
2	6.88	1.42	3.98	3.75	4.64	1.31	1.60	1.45	2.34	0.44	0.99	0.61
3	7.74	2.86	3.43	1.92	3.60	1.50	0.95	1.06	2.03	0.45	0.53	0.63
4	6.71	1.50	2.08	1.74	2.34	0.80	0.91	0.96	1.23	0.43	0.50	0.48
5	9.63	1.31	2.07	1.23	1.76	0.81	0.65	0.90	1.40	0.46	0.39	0.46
6	4.47	1.76	1.29	1.01	2.77	1.02	0.86	0.84	1.53	0.55	0.47	0.43
7	2.30	1.74	1.72	0.74	2.20	0.59	0.89	0.83	1.24	0.28	0.39	0.39
8	3.84	0.84	0.57	0.78	1.11	0.51	0.44	0.52	0.60	0.23	0.21	0.27
9	3.27	0.78	0.51	0.47	1.49	0.43	0.56	0.48	0.78	0.22	0.28	0.21
10	3.38	0.31	1.60	0.76	1.32	0.37	0.66	0.58	0.72	0.18	0.32	0.26

grid scenario	Panel 1b: investment spike size mean and maximum deviation across shocks				Panel 2b: capital stock mean and maximum deviation across shocks				Panel 3b: welfare mean and maximum deviation across shocks			
	VFI deviation:		VFI-INT dev.:		VFI deviation:		VFI-INT dev.:		VFI deviation:		VFI-INT dev.:	
	mean	max	mean	max	mean	max	mean	max	mean	max	mean	max
1	9.03	17.41	5.35	15.72	4.76	12.52	1.78	4.71	2.61	6.33	1.12	3.25
2	6.88	18.52	3.75	13.54	4.64	9.98	1.45	3.39	2.34	4.32	0.61	1.09
3	7.74	15.15	1.92	11.16	3.60	6.11	1.06	2.35	2.03	3.15	0.63	1.44
4	6.71	15.50	1.74	4.02	2.34	5.69	0.96	2.41	1.23	2.70	0.48	1.13
5	9.63	15.35	1.23	2.66	1.76	4.21	0.90	2.16	1.40	2.18	0.46	0.89
6	4.47	16.09	1.01	3.67	2.77	7.81	0.84	1.64	1.53	4.83	0.43	0.99
7	2.30	7.11	0.74	1.92	2.20	4.12	0.83	1.57	1.24	1.96	0.39	0.66
8	3.84	12.00	0.78	3.24	1.11	3.38	0.52	1.79	0.60	1.81	0.27	0.68
9	3.27	13.16	0.47	1.57	1.49	2.86	0.48	1.29	0.78	2.01	0.21	0.42
10	3.38	11.96	0.76	3.87	1.32	2.89	0.58	1.18	0.72	1.66	0.26	0.66

VFI: Value Function Iteration, VFI-INT: VFI with 35 interpolation points on each side of a capital grid point, EGM: Endogenous Grid Method, FEM: Finite Element Method. We simulate the (S, s) behavior of the model for each productivity shock value. We then average the absolute % difference for each shock from the true statistics across simulations. Panel 1(a) shows these differences for the average investment spike size. Panels 2(a) and 3(a) show the corresponding differences for the average capital stock and welfare, respectively. To capture the variation across shocks, Panel 1(b) shows the maximum deviations of the investment spike size along with the mean deviations. Fig 2(b) and 3(b) show the corresponding maximum and mean differences for the capital stock and welfare, respectively.

For each grid scenario 1-10, Table 3 shows in panel (1a) the average (across all shocks) absolute percentage deviation of investment spike sizes implied by the four approximation methods from the true solution. This table shows that the spike size deviation of FEM, EGM and VFI-INT from the true solution is for most grid scenarios very similar across these methods. For scenarios with a relatively coarse grid, investment spike sizes are noticeably different from the true solution. The relatively large distance between grid points prevents a precise determination of the threshold which is also reflected in the inaccuracy statistics around the threshold reported in Table 2 (columns 7 and 8). However, for somewhat finer grids (scenarios 4-10) the mean absolute deviation from the true solution is rather small, between 0.31% and 2.07% indicating a very precise approximation of the size of the jump discontinuity. This is very much in contrast to comparable statistics for VFI (solid blue). The average absolute investment spike size deviation from the true solution using VFI is even for grid scenarios 4-10 rather large – between 2.30%-9.63%. Also the deviation in comparison to the three other approximation methods is economically significant, e.g. for scenario 5, VFI implies an average absolute investment spike size deviation that is approximately 8 percentage points above the ones implied by VFI-INT, FEM and EGM. Overall, this table clearly shows that VFI produces, even for very fine grids, substantial and economically significant deviations from the true solution and the other methods.

Panels (2a) and (3a) show comparable statistics for the average absolute deviation of the mean of capital and welfare from the true solution. The statistics on the mean of capital statistic gives an indication about the precision of the approximated location and size of the jump discontinuity.[34] The results for the mean of capital and welfare are qualitatively very similar to the case considering investment spike size: FEM, EGM and VFI-INT deliver very similar statistics that are relatively close to the true solution. In contrast, average absolute capital mean deviations and welfare deviations generated using VFI are rather far away from the true solution and are economically rather different to the ones delivered by the other three approximation methods. Grid scenario 2 for example implies an average absolute deviation from the capital stock (welfare) that is more than 3 (1.3) percentage points higher

[34]Given the (S, s) behavior of capital, the mean of capital provides information on both the size of investment spikes and the size of the capital stock when investment is undertaken.

than for the other three methods. We see it as a success of VFI-INT that the percentage deviations implied by this method are in the range of the ones implied by FEM and EGM, despite the use of a finite number of interpolation points. This finding is especially useful when taking into account implementation and computation time of the methods, which will be discussed below.[35]

Importantly, there are considerable differences in the deviations across shocks that are not evident from the means reported in panels (1a), (2a) and (3a). To highlight these differences, we show in panels (1b), (2b) and (3b) the absolute average deviation for the three statistics of VFI and VFI-INT and the corresponding maximum deviations across all shocks. For most grid scenarios with finer grids, even the maximum deviation of VFI-INT is below the average deviation of VFI. Considering the maximum deviation across shocks further highlights the economic significance between methods. For example, VFI implies for grid scenario 6 that a firm's investment spike size deviates up to 16.09% from the true size. In contrast VFI-INT only implies a maximum deviation of 3.67%. VFI implies for the same scenario for a firm's mean capital stock a deviation up to 7.81% from the true statistics, while this value for VFI-INT is only 1.64%.

While we use a specific model to exemplify the problems of VFI to accurately identify a jump discontinuity in policy functions, note from the exposition above that these problems will be present in any application in which the location of discontinuities are determined by the intersection of option values. The economic significance of the differences between VFI and the other three methods clearly depends on specific the model and the parameterization. Note that our parameterization is relatively conservative. Differences between VFI and the other two methods would be even more pronounced for other, widely used, parameter values in the literature. For example, our value for the parameter determining the returns of capital,

[35] The interpolation in VFI-INT can be vectorized which implies that additional interpolation points come at very low additional costs in terms of computation time. Results with different numbers of interpolation points are available upon request.

$\alpha = 0.592$, is at the upper bound of used values. Lower values for α emphasize the problems of VFI to identify the threshold as it leads to flatter option values V^i and V^a.[36] Appendix A.6 shows that the results described above are robust for various alternative parameterizations. Appendix A.8 further highlights the generality of the results shown above. In particular, it shows the described superiority of VFI-INT, FEM and EGM over VFI in the ability to approximate models with jump discontinuities also holds in the context of a general equilibrium model.

We have shown that the reliance on VFI to approximate the well-known and widely used model of Cooper and Haltiwanger (2006) can be problematic. However, our results apply more broadly as many other models with jump discontinuities in policy functions are approximated using VFI in the literature (e.g. Adda and Cooper (2000)). Our analysis shows that approximations based on VFI can be highly inaccurate unless extremely fine capital grids are used.

5.3 Method Comparison

The discussion above shows that VFI-INT, EGM and FEM can all address the problems encountered by VFI in approximating policy functions with jump discontinuities. However, there are pros and cons between these methods in terms of speed and implementation complexity. Traditionally, speed across methods is compared for a given level of accuracy as measured by average log absolute Euler equation errors. This type of comparison is provided in Table 2 where, consistent with the literature, EGM is by far the fastest method. FEM is by far the slowest as the root-finding problem is very time consuming whereas VFI and VFI-INT are of roughly comparable speed for most grid scenarios and much faster than

[36] Commonly used values are between 0.30 and 0.42, see for example Gomes (2001), Görtz and Tsoukalas (2013) and King et al. (1988).

FEM.[37] [38]

Table 4: Computing Time Across Methods (in seconds).

grid scenario	VFI grid points	VFI CPU(s)	VFI-INT grid points	VFI-INT CPU(s)	FEM grid points	FEM CPU(s)	EGM grid points	EGM CPU(s)
1	350	31.97	115	22.77	27	81.34	34	2.26
2	600	67.12	276	44.71	83	228.56	70	3.72
3	700	81.46	385	66.46	95	250.70	97	5.49
4	850	107.83	400	68.68	150	407.53	120	7.48
5	950	125.88	500	90.76	175	471.26	128	7.77
6	1050	147.91	750	157.41	250	682.25	180	13.82
7	1400	262.07	1000	236.27	330	852.61	210	20.63
8	1900	416.55	1700	570.21	550	1424.64	400	82.81
9	2300	591.61	2300	1035.58	800	2158.81	500	146.83
10	3000	1126.59	2500	1179.59	1000	2645.43	650	200.16

VFI: Value Function Iteration, VFI-INT: VFI with 35 interpolation points on each side of a capital grid point, EGM: Endogenous Grid Method, FEM: Finite Element Method. For each grid scenario methods are comparable in terms of average Euler equation errors. CPU time for FEM and VFI interpolated is reported utilising four processing units. Parallelization did not improve the performance of VFI and EGM.

It is important to note however that this type of benchmarking can be misleading for models with jump discontinuities in policy functions because, as shown above, Euler equation errors are not a sufficient measure to determine the accuracy of numerical approximations for such models. From Table 3 one can see that FEM, EGM and VFI-INT for corresponding grid scenarios provide a comparable level of precision in terms of considered statistics. This implies that the corresponding grid scenarios, and therefore also the computation times shown in Table 4, are roughly comparable. However, one can also see that much higher grid scenarios are needed for VFI to produce a similar level of precision in Table 3. For example, VFI scenario 10 produces a precision comparable to scenario 3 for the other methods, implying that VFI (1126.59 seconds) takes about four times as long as FEM (250.70 seconds). So for our model with jump discontinuities in the policy function, VFI is actually much slower

[37] We ran all programs on an Intel i7-3770 (3.4 GHz) Processor with 4 active cores and 16 GB of memory running Windows 7. As we implemented all methods using Matlab, we can directly compare running time and implementation complexity.

[38] We only discuss computation speed for the model with fixed costs, however following a suggestion by a referee, we show in Appendix A.7 that the performance ranking across methods is retained also for the case without fixed costs ($F = 0$).

than FEM when benchmarking accuracy in terms of deviations from true statistics.

In general the applicability of EGM is problem dependent and for this reason EGM is by far the most complex algorithm to implement.[39] One important limitation of EGM is that it is limited to one continuous choice variable unless it is combined with additional VFI steps. Additional control or state variables require a number of intricate extensions that are often problem specific and may necessitate the use of higher dimensional interpolation (see for example Barillas and Fernandez-Villaverde (2007), Hintermaier and Koeniger (2010), Fella (2014) and Ludwig and Schön (2013)). For models with jump discontinuities, demarcating the non-concave region adds substantial programming complexity, and using VFI-INT rather than VFI in extension to EGM is then more appropriate in light of our findings. VFI, VFI-INT and FEM are far simpler to implement and easily extend to additional state and control variables. The combination of computational speed and relatively easy implementation and adaptation make VFI-INT especially suitable for approximating models with jump discontinuities in the policy functions.

6 Conclusion

Differences across approximation methods have been extensively studied for dynamic economies where policy functions are continuous. However, the literature provides little guidance about the adequacy and accuracy of computational methods for dynamic economies where agents face non-concave problems. This paper is a first attempt to fill this gap. We highlight that for models with jump discontinuities in policy functions (i) using Value Function Iteration (VFI) is problematic as it fails to accurately identify both the location and size of jump discontinuities; and (ii) Euler equation errors are not a sufficient measure for accuracy as they do not provide indications about how well the location of the discontinu-

[39] Conventionally reported measures of complexity such as code length imply EGM (300 lines of code) is much more intricate to implement than the other methods (120 lines of code).

ity is approximated. We show that much more accurate approximations for this class of models are delivered by the Endogenous Grid Method (EGM), the Finite Element Method (FEM) and value function iteration when extended with a local interpolation step (VFI-INT). We employ a well established model of a plant where investment is subject to fixed adjustment costs to compare key statistics from simulations across methods. We show that differences between policy functions generated by VFI and the three alternative methods are economically significant. As these differences across methods cannot be identified using Euler equation errors, also the conventional speed comparisons which rely on these as a measure for benchmarking accuracy can be misleading.

References

Adda, J. and Cooper, R. (2000). Balladurette and juppette: A discrete analysis of scrapping subsidies. *Journal of Political Economy*, 108:778–806.

Arellano, C. (2008). Default risk and income fluctuations in emerging economies. *American Economic Review*, 98(3):690–712.

Arellano, C., Maliar, L., Maliar, S., and Tsyrennikov, V. (2016). Envelope condition method with an application to default risk models. *Journal of Economic Dynamics and Control*, 69(C):436–459.

Arellano, C. and Ramanarayanan, A. (2012). Default and the Maturity Structure in Sovereign Bonds. *Journal of Political Economy*, 120(2):187–232.

Aruoba, S. B., Fernandez-Villaverde, J., and Rubio-Ramirez, J. F. (2006). Comparing solution methods for dynamic equilibrium economies. *Journal of Economic Dynamics and Control*, 30(12):2477–2508.

Bajari, P., Chan, P., Krueger, D., and Miller, D. (2013). A dynamic model of housing demand: Estimation and policy implications. *International Economic Review*, 54(2):409–442.

Barillas, F. and Fernandez-Villaverde, J. (2007). A generalization of the endogenous grid method. *Journal of Economic Dynamics and Control*, 31(8):2698 – 2712.

Bayer, C. (2006). Investment dynamics with fixed capital adjustment cost and capital market imperfections. *Journal of Monetary Economics*, 53(8):1909–1947.

Bloom, N. (2009). The impact of uncertainty shocks. *Econometrica*, 77(3):623–685.

Caballero, R. J., Engel, E. M. R. A., Haltiwanger, J. C., Woodford, M., and Hall, R. E. (1995). Plant-level adjustment and aggregate investment dynamics. *Brookings Papers on Economic Activity*, 1995(2):1–54.

Carroll, C. D. (2006). The method of endogenous gridpoints for solving dynamic stochastic optimization problems. *Economics Letters*, 91(3):312–320.

Clausen, A. and Strub, C. (2012). Envelope theorems for non-smooth and non-concave optimization. *Department of Economics - University of Zurich: ECON - Working Papers*, (062).

Cooper, R., Haltiwanger, J., and Power, L. (1999). Machine replacement and the business cycle: Lumps and bumps. *American Economic Review*, 89(4):921–946.

Cooper, R. W. and Haltiwanger, J. (2006). On the nature of capital adjustment costs. *Review of Economic Studies*, 73(3):611–633.

Doms, M. E. and Dunne, T. (1998). Capital adjustment patterns in manufacturing plants. *Review of Economic Dynamics*, 1(2):409–429.

Fella, G. (2014). A generalized endogenous grid method for non-smooth and non-concave problems. *Review of Economic Dynamics*, 17(2):329–344.

Gomes, J. F. (2001). Financing Investment. *American Economic Review*, 91(5):1263–1285.

Görtz, C. and Tsoukalas, J. D. (2013). Learning, Capital Embodied Technology and Aggregate Fluctuations. *Review of Economic Dynamics*, 16(4):708–723.

Heer, B. and Maußner, A. (2008). Computation Of Business Cycle Models: A Comparison Of Numerical Methods. *Macroeconomic Dynamics*, 12(05):641–663.

Hennessy, C. and Whited, T. (2005). Debt dynamics. *Journal of Finance*, 60(3):1129–1165.

Hintermaier, T. and Koeniger, W. (2010). The method of endogenous gridpoints with occasionally binding constraints among endogenous variables. *Journal of Economic Dynamics and Control*, 34(10):2074–2088.

Jose Luengo-Prado, M. (2006). Durables, nondurables, down payments and consumption excesses. *Journal of Monetary Economics*, 53(7):1509–1539.

Khan, A. and Ravikumar, B. (2002). Costly technology adoption and capital accumulation. *Review of Economic Dynamics*, 5(2):489 – 502.

Khan, A. and Thomas, J. K. (2008). Idiosyncratic shocks and the role of nonconvexities in plant and aggregate investment dynamics. *Econometrica*, 76(2):395–436.

King, R. G., Plosser, C. I., and Rebelo, S. T. (1988). Production, growth and business cycles : I. The basic neoclassical model. *Journal of Monetary Economics*, 21(2-3):195–232.

Kopecky, K. and Suen, R. (2010). Finite State Markov-chain Approximations to Highly Persistent Processes. *Review of Economic Dynamics*, 13(3):701–714.

Ludwig, A. and Schön, M. (2013). Endogenous grids in higher dimensions: Delaunay interpolation and hybrid methods. MEA discussion paper series 13274, Munich Center for the Economics of Aging (MEA) at the Max Planck Institute for Social Law and Social Policy.

McGrattan, E. R. (1996). Solving the stochastic growth model with a finite element method. *Journal of Economic Dynamics and Control*, 20(1-3):19–42.

Nilsen, O. A. and Schiantarelli, F. (2003). Zeros and lumps in investment: Empirical evidence on irreversibilities and nonconvexities. *The Review of Economics and Statistics*, 85(4):1021–1037.

Power, L. (1998). The missing link: Technology, investment, and productivity. *The Review of Economics and Statistics*, 80(2):300–313.

Rouwenhorst, K. (1995). Asset pricing implications of equilibrium business cycle models. In Cooley, T., editor, *Frontiers of Business Cycle Research*, pages 294–330.

Santos, M. S. (2000). Accuracy of numerical solutions using the euler equation residuals. *Econometrica*, 68(6):1377–1402.

Santos, M. S. and Peralta-Alva, A. (2012). Analysis of numerical errors. Working Papers 2012-6, University of Miami, Department of Economics.

Tauchen, G. (1986). Finite state markov-chain approximations to univariate and vector autoregressions. *Economics Letters*, 20(2):177–181.

Tsoukalas, J., Tsoukas, S., and Guariglia, A. (2016). To what extent are savings-cash flow sensitivities informative to test for capital market imperfections? *Review of Finance*.

Vayanos, D. (1998). Transaction costs and asset prices: A dynamic equilibrium analysis. *Review of Financial Studies*, 11(1):1–58.

Wang, P. and Wen, Y. (2012). Hayashi meets kiyotaki and moore: A theory of capital adjustment costs. *Review of Economic Dynamics*, 15(2):207 – 225.

Whited, T. M. (2006). External finance constraints and the intertemporal pattern of intermittent investment. *Journal of Financial Economics*, 81(3):467–502.