A General and Intuitive Envelope Theorem[*]

Andrew Clausen[†]
University of Edinburgh

Carlo Strub[‡]
University of St. Gallen

8th April 2016

Abstract

Previous envelope theorems establish differentiability of value functions in convex settings. Our envelope theorem applies to all functions whose derivatives appear in first-order conditions, and in non-convex settings. For example, in Stackelberg games, the leader's first-order condition involves the derivative of the follower's policy. Similarly, we differentiate (i) the borrower's value function and default cut-off policy function in an unsecured credit economy, (ii) the firm's value function in a capital adjustment problem with fixed costs, and (iii) the households' value functions in insurance arrangements with indivisible goods. Our theorem accommodates optimization problems involving discrete choices, infinite horizon stochastic dynamic programming, and Inada conditions.

Keywords: First-order conditions, policy functions, discrete choice, Inada conditions, dynamic programming, reverse calculus.

1 Introduction

A fundamental insight of economics is that optimal choices occur where marginal benefit equals marginal cost. In simple economies, both sides of this first-order condition are exogenous, and can be assumed to exist. In recursive macroeconomies,

[*]An earlier version of this paper circulated under the title "Envelope Theorems for Non-Smooth and Non-Concave Optimization." We are very grateful to Michael Elsby, Philipp Kircher, Dirk Krueger, Tzuo Hann Law, George Mailath, Jochen Mankart, Steven Matthews, Nirav Mehta, Guido Menzio, Leonard Mirman, Kurt Mitman, John H. Moore, Georg Nöldeke, Andrew Postlewaite, Kevin Reffett, Felipe Saffie, József Sákovics, Ilya Segal, Shouyong Shi, Philip Short, Jonathan Thomas, Xi Weng, Tim Worrall, and Mark Wright for fruitful discussions.

[†]andrew.clausen@ed.ac.uk

[‡]cs@carlostrub.ch

the marginal benefit of preparing for the future is endogenous, and envelope theorems have established its existence in well-behaved convex settings. However, there are many important economic problems where it is unknown whether first-order conditions apply.

Threats to first-order conditions. *Jumps* can arise in objective functions even if all of the model primitives are continuous. For example, consider the Stackelberg duopoly game, which we study in detail in Section 2. The follower's policy function appears inside the leader's objective function. But strong conditions are required to ensure that the follower's policy function is continuous. If the follower's policy function is discontinuous, then the leader's objective function is not continuous, let alone differentiable.

Kinks can arise in objective functions even if all model primitives are differentiable. This occurs when a continuous choice is taken along side a discrete choice. For example, consider a Stackelberg leader who can build his factory in China or Europe. Assume each location has a differentiable cost function. Despite this assumption, the firm's overall cost function has a kink at the quantity where both locations are equally costly.

Hidden jumps and kinks arise when the objective is differentiable, but ingredients such as benefit and cost are not. For example, suppose that the Stackelberg leader does not know the follower's cost curve. He assigns probabilities to two different follower cost curves, and hence to two different leader benefit functions. Even if the expected benefit of selling output is differentiable, this does *not* imply that the ex post benefit curves are differentiable. They might have kinks or jumps that cancel each other out. In this case, it is impossible to write a meaningful first-order condition.

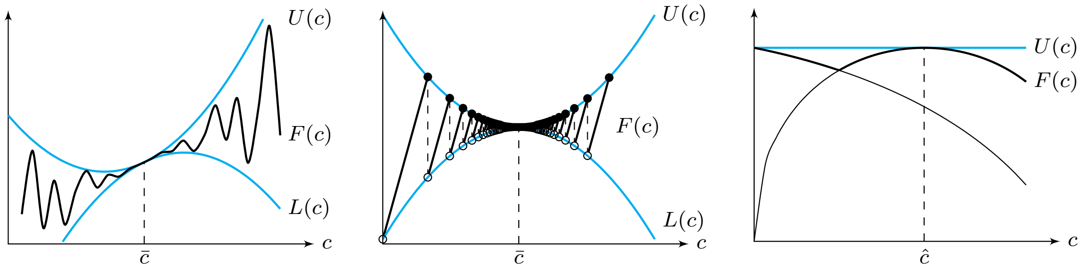
Boundary problems arise when decision makers prefer to make boundary choices, such as exhausting capacity constraints. Since first-order conditions only apply to interior solutions, economists often steer decision makers to the interior by imposing Inada conditions. This is problematic as the relevant envelope theorems depend on placing uniform bounds on derivatives. In other words, economists wishing to apply first-order conditions have an uncomfortable choice between (i) assuming Inada conditions hold to ensure that all solutions are interior, or (ii) assuming that Inada conditions do not hold, to ensure that derivatives exist.

These threats are common place in important economic problems. For example, all four threats arise in the unsecured credit market model of Arellano (2008), which we study in Section 5.1.[1] First, the borrower's future default policy appears in his objective, because it determines default risk and hence interest rate pay-

[1] Below, we also discuss related work including Eaton and Gersovitz (1981), Aguiar and Gopinath (2006), Hatchondo and Martinez (2009), and Arellano and Ramanarayanan (2012).

ments. There is no a priori reason why his policy would be differentiable. Second, the borrower has a discrete choice – whether to honour or default on debts owed – leading to kinks in his value function. Third, even if the objective is differentiable, the default policy and value function might have jumps or kinks that cancel each other out. Fourth, Arellano focuses on a utility function that satisfies the Inada conditions. These four features of Arellano's model pose difficulties to applying any of the existing envelope theorems.

Techniques. We devise three new techniques for addressing these threats to first-order conditions. Our *Differentiable Sandwich Lemma* synthesises various sandwich arguments deployed in previous envelope theorems. It establishes that a function F is differentiable at a point $\bar{c}$ if it is sandwiched between two differentiable functions U and L, as depicted in Figure 1. Specifically, the lemma applies if the two functions, which we call differentiable upper and lower support functions, satisfy (i) $U(\bar{c}) = F(\bar{c}) = L(\bar{c})$, (ii) $U(c) \geq F(c) \geq L(c)$ for all c, and (iii) L and U are differentiable at $\bar{c}$. Our lemma generalises Benveniste and Scheinkman (1979, Lemma 1) by dropping all convexity requirements.



(a) F is differentiable at $\bar{c}$ (b) F is differentiable at $\bar{c}$ (c) Upper support at a max

Figure 1: Differentiable Sandwich Lemma

The lemma is well-suited to studying optimal choices. Suppose that the decision maker must make a continuous choice $c \in \mathbb{R}$ followed by a discrete choice $d \in D$. Assume that his utility function $v_d(c)$ is differentiable in c for each discrete choice d. Let $F(c) = \max_{d \in D} v_d(c)$ be the value after choosing c. Notice that at an optimal choice $(\hat{c}, \hat{d})$, the value function is sandwiched between the horizontal line $U(c) = F(\hat{c})$ and $v_{\hat{d}}$, as depicted in Figure 1c. Therefore, F is differentiable at $\hat{c}$. Milgrom and Segal (2002, Corollary 2) previously drew this conclusion for the special case that $\{v_d\}_{d \in D}$ is equi-differentiable and has uniformly bounded derivatives. Their redundant conditions conflict with Inada conditions. This means that the Differentiable Sandwich Lemma is applicable to problems with discrete choices and Inada conditions for the first time.

Our second and most important innovation, *Reverse Calculus*, is the opposite of normal calculus. Whereas normal calculus establishes that $H(c) = F(c) + G(c)$ is differentiable if F and G are differentiable, reverse calculus establishes that F and G are differentiable if H is differentiable. The main requirement for reverse calculus is that each ingredient function must have an appropriate differentiable support function. For example, if $H(c) = F(c) + G(c)$, then we require F and G have differentiable lower support functions f and g at $\bar{c}$, depicted in Figure 2. Under these conditions, F is sandwiched between f and $H - g$, and is therefore differentiable.

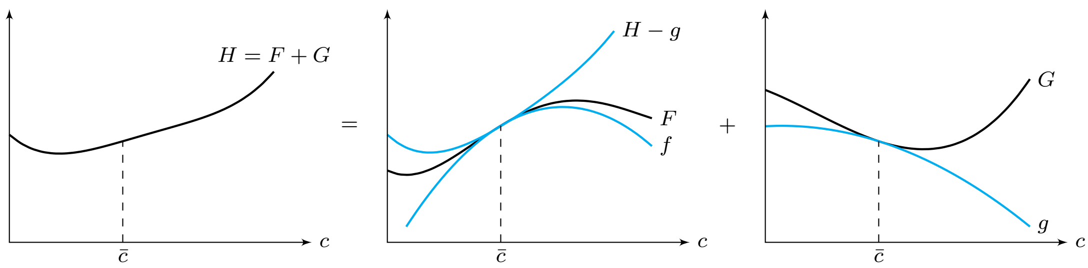


Figure 2: Reverse calculus: F is differentiable at $\bar{c}$.

Reverse calculus addresses problems whose objectives involve policy functions and/or expectations over a family of value functions. As discussed above, it is insufficient to establish that the objective function (e.g. H) is differentiable. Meaningful first-order conditions require us to establish that all of the relevant ingredient functions (e.g. F and G) are differentiable. We develop a reverse calculus for many standard operations, including addition, multiplication, function composition and upper envelopes.

Our third innovation addresses a remaining problem: finding an appropriate differentiable support function for policies. Unlike our previous innovations, it applies only to optimal stopping rules such as when to default on debt obligations. Benveniste and Scheinkman's (1979) proof involved constructing a value function of a "lazy" decision maker that uses a completely unresponsive policy rule. Since this lazy decision maker makes weakly inferior choices, his value function is a lower support function. Moreover, since the unresponsive policy is constant, the lazy value function merely traces out the shape of the differentiable utility function. Benveniste and Scheinkman conclude that the lazy value function is a differentiable lower support function. We apply the same idea to optimal stopping rules. Specifically, we construct a *lazy stopping rule* derived from underestimating the surplus from continuing. Lazy stopping rules terminate prematurely, and hence provide a differentiable support function.

Contribution. These techniques are complementary, and allow us to prove the following theorem: if all ingredient functions of an optimization problem have appropriate differentiable support functions, then at any optimal choice, (i) the objective is differentiable, and (ii) all of the ingredients are differentiable. This theorem applies even when there are discrete choices, the primitives involve Inada conditions, and many endogenous ingredient functions are combined in expectations, budget constraints, or incentive constraints.

This allows us to resolve several open problems. We re-examine three economies in which previous authors have applied first-order conditions, where the correctness of these conditions was until now an open question. Specifically, we establish that first-order conditions do hold (i) in unsecured credit markets with endogenous default probabilities, (ii) in capital markets with fixed costs of adjustment,[2] and (iii) in informal insurance arrangements with indivisible choices.[3]

Outline. Section 2 lays out our recipe for deriving first-order conditions using a Stackelberg duopoly as a running example. Section 3 formally specifies and proves the lemmas, and combines them into an abstract envelope theorem. It also develops a novel proof of the Benveniste and Scheinkman (1979) envelope theorem. In Section 4 we compare our techniques to previous work. In Section 5, we apply them to the open questions listed above, and Section 6 concludes. The appendix connects differentiable sandwiches with Fréchet subderivatives.

2 Illustration

This section develops a recipe for deriving first-order conditions through a series of illustrated examples. The examples are all Stackelberg duopoly games in which a leader's first-order condition involves the derivative of a follower's policy function. Other related problems are explored by Kydland and Prescott (1977) and Ljungqvist and Sargent (2012, Chapter 19). The textbook analysis requires strong convexity and twice-differentiability assumptions, which we relax by introducing capacity constraints. The follower's policy has a kink where he switches from being constrained to unconstrained. We establish that the leader steers the follower away from these kinks.

[2] Below, we discuss the work of Harrison, Sellke and Taylor (1983), Caballero and Engel (1999), Cooper and Haltiwanger (2006), Gertler and Leahy (2008), Khan and Thomas (2008b) and Elsby and Michaels (2014).

[3] Below, we discuss the work of Thomas and Worrall (1988, 1990), Kocherlakota (1996), Ligon, Thomas and Worrall (2002), Koeppl (2006), Rincón-Zapatero and Santos (2009), and Morten (2013).

The first example illustrates how to apply the differentiable sandwich lemma to obtain first-order conditions when there is only one policy function. In the second example, there are two policy functions. We apply reverse calculus to establish that the follower steers the leader away from the kinks in both policy functions. While reverse calculus applies generally to optimisation problems, the third example shows that it generically fails elsewhere.

Textbook Stackelberg Competition. To fix notation, we review the textbook analysis (e.g. Varian (1992, Section 16.6)) of a Stackelberg duopoly. A leader and a follower choose their output levels, y_1 and y_2 sequentially, which costs them $C(y_1)$ and $C(y_2)$. The output is sold at the market price, $P(y_1+y_2)$. Firm i earns profits $\pi_i(y_1, y_2) = y_i P(y_1 + y_2) - C(y_i)$. The follower chooses $y_2 = f(y_1)$ by solving

$$f(y_1) = \arg\max_{y_2} \pi_2(y_1, y_2) \tag{1}$$

and the leader chooses the y_1 that maximises his objective

$$\phi_1(y_1) = \pi_1(y_1, f(y_1)). \tag{2}$$

The textbook first-order conditions for the follower and leader are

$$P(y_1 + y_2) + P'(y_1 + y_2)y_2 = C'(y_2) \text{ and} \tag{3}$$

$$P(y_1 + f(y_1)) + P'(y_1 + f(y_1))(1 + f'(y_1))y_1 = C'(y_1). \tag{4}$$

The derivatives of P, f, and C appear in (4). The demand function P and cost function C are exogenous, so we can assume they are differentiable. However, we can not assume that the follower chooses a differentiable policy f.

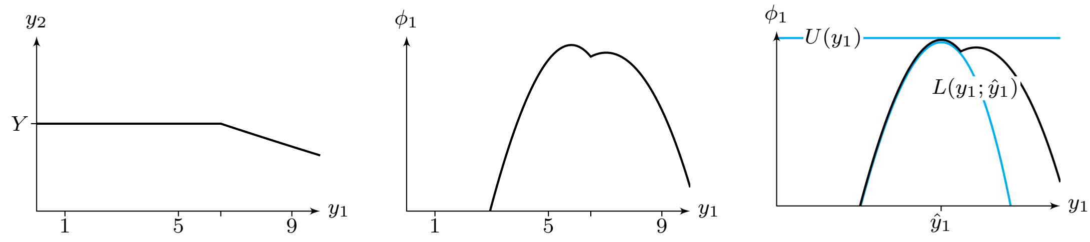
The textbook solution is to assume the cost and demand functions are strictly convex/concave and twice differentiable. Under these assumptions, (3) implicitly defines the follower's policy function, which is differentiable by the implicit function theorem. Therefore, the leader's first-order condition (4) holds at her optimal choices.

Example 1: Stackelberg with Capacity Constraints. Suppose the follower has a capacity constraint Y. The follower's cost function is

$$C(y) = \begin{cases} c(y) & \text{if } y \le Y, \\ \infty & \text{if } y > Y, \end{cases} \tag{5}$$

where $c(\cdot)$ is twice differentiable and strictly convex.

The follower's best response has an upward kink, depicted in Figure 3a, where his capacity constraint transitions from binding to non-binding. This translates



(a) Follower's policy (b) Leader's profit (c) Differentiable sandwich

Figure 3: Stackelberg with a capacity constrained follower

into a downward kink in the leader's objective, depicted in Figure 3b. Would the leader ever choose this kink, thus invalidating the first-order condition (4)?

The answer is no. To prove this, we will construct a differentiable sandwich around the leader's objective $\phi_1(\cdot)$ at the optimal choice $\hat{y}_1$. First, we construct the bottom half of the sandwich. Since the follower's policy (depicted in Figure 3a) is the lower envelope of two differentiable functions, it has a differentiable upper support function $F(\cdot)$ at $\hat{y}_1$. The follower's policy enters the leader's objective through the downward sloping demand curve P. This means that $L(y_1) = \pi_1(y_1, F(y_1))$ is a lower support function for the leader's objective at $\hat{y}_1$.

Second, we construct the top half of the sandwich as the constant function $U(y_1) = \phi(\hat{y}_1)$.

These support functions form a sandwich, illustrated in Figure 3c. By the differentiable sandwich lemma, the leader's objective is differentiable at $\hat{y}_1$, where it satisfies the first-order condition

$$\phi_1'(\hat{y}_1) = U'(\hat{y}_1) = 0. \tag{6}$$

However, we have not yet established (4), which is a more useful first-order condition. In particular, we have not yet determined whether f is differentiable at $\hat{y}_1$. This is a reverse calculus problem: we have established that the left side of (2) is differentiable, and we would now like to infer that the policy f on the right side is differentiable. In this example, there is a simple solution. Since (2) implicitly defines f near $\hat{y}_1$, the implicit function theorem implies f is differentiable at $\hat{y}_1$. We conclude that the leader's first-order condition (4) holds at optimal choices.

This example illustrates how the differentiable sandwich lemma and a simple form of reverse calculus can be applied to establish first-order conditions. The main task was constructing the bottom half of the sandwich. We constructed a differentiable upper support function for the follower's policy, which we used to construct a differentiable lower support function for the leader's objective. Since there was only one endogenous function, the reverse calculus step only required the implicit function theorem.

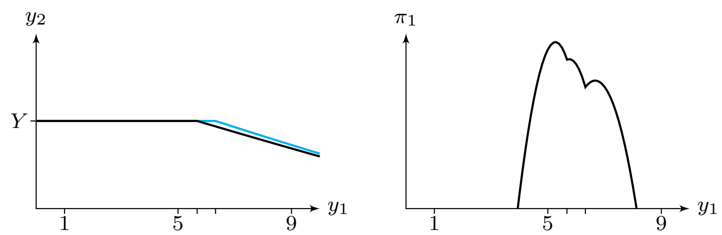
Example 2: Stochastic Stackelberg with capacity constraints. The previous example only required a simple reverse calculus step, because there was only one policy function. We now consider a problem with two policy functions. We extend Example 1 by assuming that the follower has privately known costs. Specifically, the leader assigns probabilities p_A and p_B to the follower's production cost being $C_A(\cdot)$ or $C_B(\cdot)$, respectively. The follower now has two policies f_A and f_B, one for each cost function. The leader's problem is to choose output y_1 to maximise her expected profit

$$\phi_1(y_1) = \sum_{z \in \{A,B\}} p_z y_1 P(y_1 + f_z(y_1)) - C(y_1). \tag{7}$$

We would like to determine whether her first-order condition

$$\sum_{z \in \{A,B\}} p_z \left\{ P(y_1 + f_z(y_1)) + y_1 P'(y_1 + f_z(y_1))(1 + f_z'(y_1) \right\} - C'(y_1) = 0 \tag{8}$$

holds at her optimal choice $\hat{y}_1$. As before, the follower's policy functions have one kink each, depicted in Figure 4a. The corresponding leader objective function (depicted in Figure 4b) inherits both kinks, but neither kink is an optimal choice. Does the leader always steer the follower away from the kinks in his policy functions?



(a) Follower's policies (b) Leader's profit

Figure 4: Stackelberg follower with privately known cost

As in the previous example, we can construct a differentiable upper support function $F_z(\cdot)$ for each policy $f_z(\cdot)$ at $\hat{y}_1$. The same method as before establishes that the first-order condition $\phi_1'(\hat{y}_1) = 0$ holds.

We show that the follower's policy functions $f_z(\cdot)$ are differentiable at $\hat{y}_1$. In the previous example, this was a straightforward application of the implicit function theorem. However, one equation can not implicitly define two policy functions. Instead, we apply our reverse calculus summation rule to

$$\sum_{z \in \{A,B\}} p_z y_1 P(y_1 + f_z(y_1)).$$

Since this sum is differentiable, and each term has a differentiable lower support function $L_z(y_1) = p_z y_1 P(y_1 + F_z(y_1))$, the rule implies that each term is differentiable at $\hat{y}_1$. Therefore, both policy functions are differentiable at optimal choices, and the first-order condition (8) holds.

When there are two or more policy functions, the implicit function theorem can not be applied to establish they are differentiable. This example showed how reverse calculus applies in these situations. In fact, we did not need to impose any additional conditions. This lead to the following recipe:

(i) Construct a differentiable lower support function for the objective, by finding appropriate support functions for the ingredients.

(ii) A constant upper support function exists at optimal choices.

(iii) By the differentiable sandwich lemma, the objective's derivative exists and is zero.

(iv) By reverse calculus, all derivatives in the first-order condition exist.

Example 3: Stackelberg leader's value. In the first two examples, we applied reverse calculus to show that the leader chooses differentiable points of the follower's policies. In this example, we illustrate how reverse calculus might fail. We focus our attention on the follower's value function,[4]

$$V_2(y_1) = f(y_1)P(y_1 + f(y_1)) - C(f(y_1)). \tag{9}$$

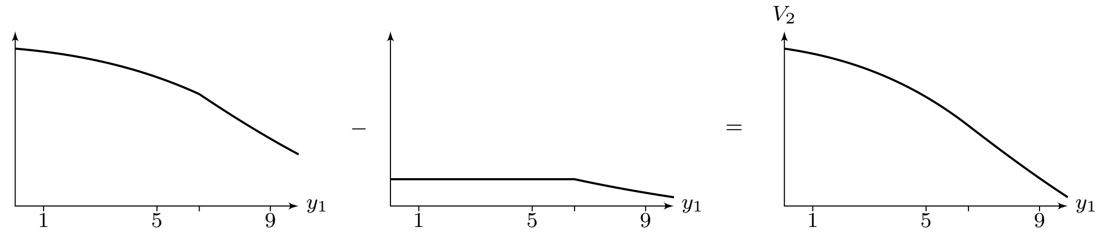
It is straightforward to show that V_2 is a concave function. Therefore, the Benveniste and Scheinkman (1979) envelope theorem (which we prove in Section 4.1) implies that V_2 is globally differentiable. It is tempting to deduce that the follower's policy, f, which appears twice on the right side of (9), must also be globally differentiable. However, in Figure 3a, we saw that f is *not* globally differentiable!

Why does reverse calculus not apply here? The right side of (9) is the difference between revenue and cost, which have identical kinks that cancel each other out. This cancellation is depicted in Figure 5. The relevant reverse calculus rule is the implicit function theorem, because the policy function $y_2 = f(y_1)$ is implicitly defined by the equation $\psi(y_1, y_2) \equiv V_2(y_1) - y_2 P(y_1 + y_2) + C(y_2) = 0$. The implicit function theorem requires that $\frac{\partial}{\partial y_2}\psi(y_1, y_2)$ be non-zero at $(y_1, f(y_1))$. But it is *always* zero – this is the follower's first-order condition for choosing y_2.

Kink cancellation is not special to value functions in Stackelberg games. It is a property of generic concave dynamic programming problems of the form

$$\begin{aligned} V(a) &= \max_{a'} u(a, a') + \beta V(a') &(10)\\ &= u(a, f(a)) + \beta V(f(a)), &(11) \end{aligned}$$

[4] Value functions of other players arise in problems with promise-keeping constraints.



(a) Follower's optimal revenue, $f(y_1)P(y_1 + f(y_1))$

(b) Follower's optimal cost, $C(f(y_1))$

(c) Follower's value function

Figure 5: Kinks exactly cancel, endogenously

where u is concave and once (but not twice) differentiable, and f is the policy function. Again, V is globally once (but not twice) differentiable. Santos (1991) showed that f is only differentiable where V is twice differentiable. We conclude that the left side of (11) is differentiable, but that the two terms on the right side are not; they contain kinks that cancel each other out where V is not twice differentiable.

This is a cautionary example. When the conditions of reverse calculus are not met, kinks in ingredient functions might be hidden by kink cancellation. This occurs in generic concave dynamic programming problems.

3 Envelope Theorem

This section presents the proofs in full generality of the differentiable sandwich lemma and of reverse calculus.

3.1 Differentiable Sandwich Lemma

Before stating the lemma, we need to be precise about what a derivative is. Since we would like to accommodate many continuous choices (such as asset portfolio choices), we use the standard multidimensional definition of differentiability. This definition ensures that the chain rule and other calculus identities are valid.

Definition 1. A function $F : C \to \mathbb{R}$ with domain $C \subseteq \mathbb{R}^n$ is **differentiable** at $c \in \text{int}(C)$ if there is some row vector m with $m^\top \in \mathbb{R}^n$ such that

$$\lim_{\Delta c \to 0} \frac{F(c + \Delta c) - F(c) - m\,\Delta c}{\|\Delta c\|} = 0. \tag{12}$$

Such an m is the **derivative** of F at c, and is denoted $F'(c)$.

In fact, this definition is almost identical to the case where the domain is a subset of a Banach space $(X, \|\cdot\|)$, and our results generalize without amendment.[5]

Lemma 1 (Differentiable Sandwich Lemma)**.** *If F is differentiably sandwiched between L and U at $\bar{c}$ then F is differentiable at $\bar{c}$ with $F'(\bar{c}) = L'(\bar{c}) = U'(\bar{c})$.*

Proof. The difference function $d(c) = U(c) - L(c)$ is minimized at $\bar{c}$. Therefore, $d'(\bar{c}) = 0$ and we conclude $L'(\bar{c}) = U'(\bar{c})$.

Let $m = L'(\bar{c}) = U'(\bar{c})$. For all Δc,

$$\frac{L(\bar{c} + \Delta c) - F(\bar{c}) - m\,\Delta c}{\|\Delta c\|} \leq \frac{F(\bar{c} + \Delta c) - F(\bar{c}) - m\,\Delta c}{\|\Delta c\|} \leq \frac{U(\bar{c} + \Delta c) - F(\bar{c}) - m\,\Delta c}{\|\Delta c\|}. \quad (13)$$

Consider the limits as $\Delta c \to 0$. Since $L'(\bar{c}) = U'(\bar{c}) = m$, the limits of the first and last fractions are 0. By Gauss' Squeeze Theorem, we conclude that the limit in the middle is also 0, and hence that F is differentiable at $\bar{c}$ with $F'(\bar{c}) = m$. □

Remark 3.1. The Differentiable Sandwich Lemma also applies when $F : C \to \mathbb{R}$ is sandwiched between L and U on an open neighbourhood of $\bar{c}$.

3.2 Maximum Lemma

The Differentiable Sandwich Lemma requires us to construct upper and lower support functions. One simple construction is a horizontal line (or hyperplane) through the maximum of a function (see Figure 3c).

Lemma 2 (Maximum Lemma)**.** *Let $\phi : C \to \mathbb{R}$ be a function. If $\hat{c} \in \mathrm{int}(C)$ maximises ϕ, then $U(c) = \phi(\hat{c})$ is a differentiable upper support function of ϕ.*

3.3 Reverse Calculus

Calculus involves rules such as "if F and G are differentiable at $\bar{c}$, then $H(c) = F(c) + G(c)$ is also differentiable at $\bar{c}$." Reverse calculus rules go in the opposite direction. We provide the most important rules here, and additional rules for convex combinations and endogenous function composition in Appendix B. We also omit the corresponding rules for subtraction and division, which would involve both upper and lower support functions.

[5] In Banach spaces, the derivative m is called a "Fréchet derivative" and lies in the topological dual space $X^* = \{m : X \to \mathbb{R} \text{ such that } m \text{ is linear and continuous}\}$. For our purposes, it is unnecessary to define a topology on X^* because all limits are taken in $(X, \|\cdot\|)$ and $\mathbb{R}$.

Lemma 3 (Reverse Calculus)**.** *Suppose $F : C \to \mathbb{R}$, and $G : C \to \mathbb{R}$ have differentiable lower support functions f, and g respectively at $\bar{c}$.*

(i) If $H(c) = F(c) + G(c)$ is differentiable at $\bar{c}$, then F is differentiable at $\bar{c}$.

(ii) If $H(c) = F(c)G(c)$ is differentiable at $\bar{c}$ and $F(\bar{c}) > 0$ and $G(\bar{c}) > 0$, then F is differentiable at $\bar{c}$.

(iii) If $H(c) = \max\{F(c), G(c)\}$ is differentiable at $\bar{c}$ and $F(\bar{c}) = H(\bar{c})$, then F is differentiable at $\bar{c}$.

(iv) If $H(c) = J(F(c))$ and $J : \mathbb{R} \to \mathbb{R}$ are differentiable at $\bar{c}$ and $F(\bar{c})$ respectively with $J'(F(\bar{c})) \neq 0$, then F is differentiable at $\bar{c}$.

Proof. Let f and g be differentiable lower support functions of F and G at $\bar{c}$. For (i)–(iii), we sandwich F between f and an appropriate differentiable upper support function U and apply the Differentiable Sandwich Lemma (Lemma 1). Appropriate upper support functions are (i) $U(c) = H(c) - g(c)$, (ii) $U(c) = H(c)/g(c)$, and (iii) $U(c) = H(c)$.

For (iv), $F(c) = J^{-1}(H(c))$ is differentiable at $\bar{c}$ by the inverse function theorem and the chain rule. □

3.4 Theorem

The recipe from Section 2 applies generally. We now show that if an objective is constructed out of endogenous functions using standard mathematical operations, then those functions' derivatives may be included in first-order conditions provided that they have appropriate differentiable support functions. The notation here is quite abstract, so we explain it using Example 2.

To establish this result, we must be more precise about what it means to construct a function out of other functions. We define an envelope algebra as the set of all functions that may be constructed from a set of (potentially endogenous) functions. Our definition is recursive to accommodate the idea that once we construct a function, we can use that function to construct other functions.

Let $\mathcal{F}(C)$ be the set of functions with domain C and co-domain $\mathbb{R}$.

Definition 2. We say $\mathcal{E} \subseteq \mathcal{F}(C)$ is an **envelope algebra** if:

(i) $F + G \in \mathcal{E}$ for all $F, G \in \mathcal{E}$,

(ii) $FG \in \mathcal{E}$ for all $F, G \in \mathcal{E}$ with $F, G : C \to \mathbb{R}_{++}$,

(iii) $H(c) = \max_{G \in \mathcal{G}} G(c)$ is in $\mathcal{E}$ for all $\mathcal{G} \subseteq \mathcal{E}$ provided it is well-defined, and

(iv) $J \circ F \in \mathcal{E}$ for all $F \in \mathcal{E}$ and all differentiable $J : \mathbb{R} \to \mathbb{R}$ with $J_1 : \mathbb{R} \to \mathbb{R}_{++}$.

Definition 3. The **generated envelope algebra** $\mathcal{E}(\mathcal{F})$ is the smallest envelope algebra generated by $\mathcal{F} \subseteq \mathcal{F}(C)$ that contains $\mathcal{F}$.

In Example 2, $\mathcal{F}$ would be the negative of the two policy functions, $\{-f_A, -f_B\}$. This negation is necessary, because we are formulating our abstract theory in terms lower support functions of the endogenous functions. The generated envelope algebra, $\mathcal{E}(\mathcal{F})$ is an infinite set that includes the leader's objective function ϕ_1.

The following lemma establishes that if all of the endogenous functions $\mathcal{F}$ have differentiable lower support functions at $\bar{c}$, then so do all of the functions constructed out of them. In particular, this means that the leader's objective ϕ_1 has a differentiable lower support function.

Lemma 4. *Let $\mathcal{F} \subseteq \mathcal{F}(C)$ be a set of functions that have a differentiable lower support function at $\bar{c} \in \operatorname{int}(C)$. Then every $F \in \mathcal{E}(\mathcal{F})$ has a differentiable lower support at $\bar{c}$.*

Now, we turn our attention to applying the Reverse Calculus Lemma (Lemma 3). Our goal is to claim that the derivatives that appear in first-order conditions exist. But which derivatives appear? To answer this question, we apply the algebraic rules of calculus. We call functions whose derivatives appear in the (algebraic formula) for the derivative of the objective *active*.

Definition 4. Fix any $(\mathcal{E}, \phi, \bar{c})$ such that $\mathcal{E}$ is an envelope algebra, $\phi \in \mathcal{E}$, and $\bar{c} \in C$. We define the **active envelope set** $\mathcal{A}(\mathcal{E}, \phi, \bar{c})$ as the smallest set $\mathcal{A} \subseteq \mathcal{E}$ such that

(i) $\phi \in \mathcal{A}$.

(ii) If $F, G \in \mathcal{E}$ and $F + G \in \mathcal{A}$, then $F, G \in \mathcal{A}$.

(iii) If $F, G \in \mathcal{E}$ and $F, G : C \to \mathbb{R}_{++}$ and $FG \in \mathcal{A}$, then $F, G \in \mathcal{A}$.

(iv) If $F \in \mathcal{G} \subseteq \mathcal{E}$ and $H(c) = \sup_{G \in \mathcal{G}} G(c)$ is in $\mathcal{A}$ and $F(\bar{c}) = H(\bar{c})$, then $F \in \mathcal{A}$.

(v) If $J \circ F \in \mathcal{A}$ where $J : \mathbb{R} \to \mathbb{R}$ is differentiable and $J_1 : \mathbb{R} \to \mathbb{R}_{++}$, then $F \in \mathcal{A}$.

Finally, we can state our main result. Informally speaking, the theorem says the following. Suppose an objective function ϕ is constructed out of functions, all of which have differentiable lower support functions. Then, at any interior optimal choice, (i) all derivatives appearing in the first-order condition exist, and (ii) a first-order condition holds. In Example 2, the theorem establishes that both policy functions' derivatives exist.

Theorem 1 (Envelope Theorem)**.** *Let $\mathcal{F} \subseteq \mathcal{F}(C)$ be a set of functions that have a differentiable lower support function at $\hat{c} \in \operatorname{int}(C)$. If $\phi \in \mathcal{E}(\mathcal{F})$ and $\hat{c} \in \arg\max_{c \in C} \phi(c)$, then (i) every function in the active function set $\mathcal{A}(\mathcal{E}(\mathcal{F}), \phi, \hat{c})$ is differentiable at $\hat{c}$, and (ii) $\phi_1(\hat{c}) = 0$.*

Proof. Since $\phi \in \mathcal{E}(\mathcal{F})$ and the envelope algebra $\mathcal{E}(\mathcal{F})$ is generated from functions with differentiable lower support functions at $\hat{c}$, Lemma 4 implies that ϕ has a differentiable lower support function at $\hat{c}$. Since $\hat{c}$ maximises ϕ, Lemma 2 establishes that ϕ has a differentiable upper support function at the maximum $\hat{c}$. Therefore, ϕ is sandwiched between two differentiable functions, so Lemma 1 implies that it is differentiable at $\hat{c}$. Moreover, $\phi'(\hat{c})$ coincides with the derivative of its upper support function, which is 0.

We prove by induction that every function in the active set $\mathcal{A} = \mathcal{A}(\mathcal{E}(\mathcal{F}), \phi, \hat{c})$ is differentiable at $\hat{c}$. We set $\mathcal{A}^1 = \{\phi\}$. To construct $\mathcal{A}^{n+1}$, we examine each $H \in \mathcal{A}^n$. For each part of Lemma 3, we select appropriate functions F and G from $\mathcal{E}(\mathcal{F})$, and conclude that F is differentiable at $\hat{c}$. We do this for every possible combination of F and G, and include each such F in $\mathcal{A}^{n+1}$. We repeat this a countable number of times, and observe that $\mathcal{A} = \cup_{n=1}^{\infty} \mathcal{A}^n$. □

Theorem 1 establishes the method from Section 2 applies to a wide class of optimisation problems. While the general setting of the theorem is quite abstract, the method itself is quite intuitive. In the next section, we explore more concrete special cases of the theorem.

4 Corollaries and Related Literature

We apply our recipe to various classes of decision problems, and contrast our results to previous literature. The main advances are that we can study all types of endogenous functions (not just value functions), we accommodate non-smooth problems involving discrete choices or boundaries, and we accommodate Inada conditions that are often imposed to simplify first-order conditions.

4.1 Value Functions in Smooth Concave Problems

Benveniste and Scheinkman (1979) study *value* functions in smooth concave dynamic programming problems, but not *policy* functions. Their main theorem establishes that value functions in this setting are differentiable. The Differentiable Sandwich Lemma leads to an elementary proof of their theorem.

Problem 1. Consider the following dynamic programming problem:

$$V(c) = \sup_{c' \in \{\bar{c}' : (c, \bar{c}') \in \Gamma\}} u(c, c') + \beta V(c'), \tag{14}$$

where the domain of V is C. We assume that (i) Γ is a convex subset of $C \times C$, (ii) u is concave, and (iii) $u(\cdot, c')$ and $u(c, \cdot)$ are differentiable, respectively.

Corollary 1 (Benveniste-Scheinkman Theorem)**.** *If $\hat{c}'$ is an optimal choice at state $c \in \operatorname{int}(\{\bar{c} : (\bar{c}, \hat{c}') \in \Gamma\})$, then V is differentiable at c with $V_1(c) = u_1(c, \hat{c}')$.*

Proof. V is concave because u is concave and Γ is convex. Hence, the supporting hyperplane theorem can be applied to the hypograph of V to construct a linear upper support function U that touches V at c. We construct the differentiable lower support function $L(c) = u(c, \hat{c}') + \beta V(\hat{c}')$. Lemma 1 delivers the conclusions. □

Graduate economics textbooks such as Stokey and Lucas (1989) do not provide a self-contained proof of this corollary. Our proof is short and elementary, and therefore suitable for junior graduate students. The original proof is based on a sandwich lemma, which Benveniste and Scheinkman (1979) prove with the help of Rockafellar (1970, Theorem 25.1). (Their lemma imposes a redundant assumption, that the lower support function be concave.) Mirman and Zilcha (1975, Lemma 1) prove a one-dimensional special case using Dini derivatives rather than sandwiches.

4.2 Value Functions in Non-Smooth Problems

Milgrom and Segal (2002) study the differentiability of *value* functions and objective functions without making any topological or convexity assumptions. Our main contribution is that we also study *policy* functions. In addition, we present two generalisations of their theorems. Our first generalisation accommodates Inada conditions, which are frequently employed to ensure first-order conditions apply. We then generalise further to accommodate stochastic dynamic programming problems.

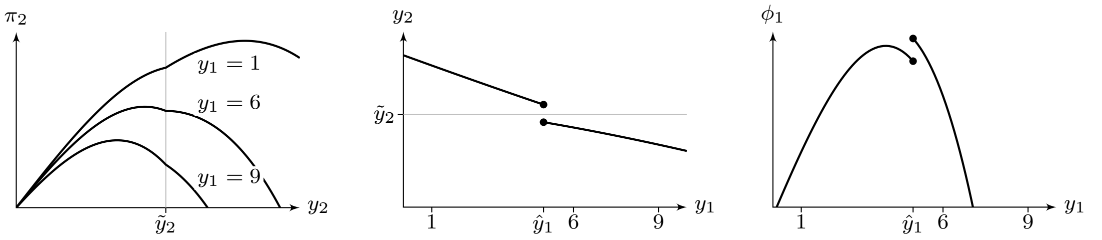
Their envelope theorems are the first to accommodate discrete choices. Specifically, they consider value functions of the form $\phi(c) = \sup_d f(c, d)$, where $\{f(\cdot, d)\}_{d \in D}$ is an arbitrary collection of differentiable functions. Here, c and d represent continuous and discrete choices, such as a quantity and factory location choice. Their Corollary 2 establishes that discrete choices only lead to downward kinks in ϕ, which decision makers would always avoid. Their result is a special case of the following corollary:

Corollary 2. *If $(\hat{c}, \hat{d}) \in \arg\max f$ then ϕ is differentiable at $\hat{c}$ with $\phi'(\hat{c}) = f_1(\hat{c}, \hat{d}) = 0$.*

Proof. The objective ϕ is differentiably sandwiched between $L(c) = f(c, \hat{d})$ and $U(c) = f(\hat{c}, \hat{d})$ at $\hat{c}$. □

We illustrate this corollary with an example, before comparing Milgrom and Segal's work with our own.

Example 4: Stackelberg with Factory Choice. Suppose the follower has access to one of two factories, which have cost functions C^1 and C^2, respectively (which are twice differentiable and strictly convex as usual). The follower's overall cost function is therefore $C(y) = \min\{C^1(y), C^2(y)\}$. The follower's profit function has a downward kink at $\tilde{y}_2$ where both factories are equally costly, as depicted in Figure 6a.



(a) Follower's profit, $\pi_2(y_1, y_2)$ (b) Follower's policy, $f(y_1)$ (c) Leader's objective function, $\phi_1(y_1) = \pi_1(y_1, f(y_1))$

Figure 6: Stackelberg when the follower has a discrete choice

Corollary 2 establishes that decision makers never choose kinks like $\tilde{y}_2$, i.e. that arise from discrete choices. This means that decision makers only choose points where their objective is differentiable. However, this does not mean that the textbook analysis of Stackelberg games generalises to this example. The follower's policy function and leader's value function, depicted in Figure 6b and Figure 6c, are discontinuous. The leader's first-order condition (4) does not hold at her optimal choice, $\hat{y}_1$.

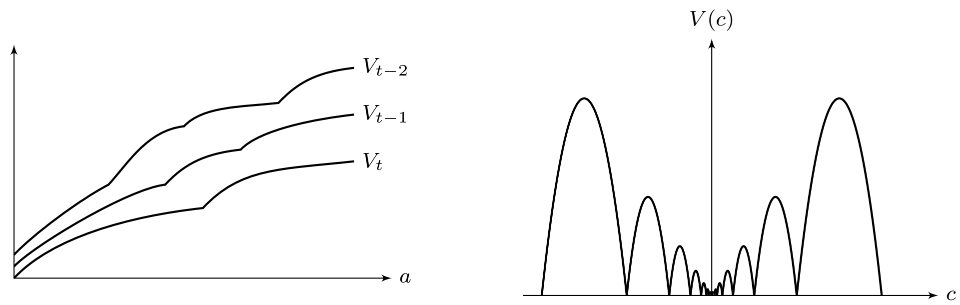
Milgrom and Segal (2002, Corollary 2) is a special case of our Corollary 2. Their version imposes redundant equidifferentiability and bounded derivative conditions. These extra conditions ensure that directional derivatives of ϕ exist globally, which their proof makes use of. Similarly, other papers make assumptions including Lipschitz continuity,[6] and supermodularity[7] to ensure the existence of directional derivatives. Our method does not require any such assumptions.

Both of their redundant conditions are problematic. The uniformly bounded derivative condition conflicts with Inada conditions. Inada conditions are imposed to ensure first-order conditions hold by directing optimal choices away from boundaries.

The equidifferentiability condition is problematic in dynamic problems, when a discrete decision is taken in every period. The number of combinations of discrete choices increases exponentially as the number of periods increases, which can lead

[6] Clarke (1975)
[7] Amir, Mirman and Perkins (1991)

the number of kinks to grow rapidly. In other words, the kinks from tomorrow's value function propagate into today's value function, as depicted in Figure 7a. In infinite horizon problems, the set of combinations of discrete choices is (uncountably) infinite. This can cause directional differentiability of value functions can fail. For example, the "bouncing ball" function depicted in Figure 7b has no directional derivatives at $c = 0$, even though it is the upper envelope of functions with uniformly bounded derivatives.[8]



(a) Backward propagation of downward kinks

(b) The bouncing ball function has no directional derivatives at $c = 0$

Figure 7

Our differentiable sandwich approach overcomes these obstacles, so we can provide the first general envelope theorem for dynamic programming problems involving discrete choices.

Problem 2. Consider the following stochastic dynamic programming problem:

$$V(c,d,\theta) = \sup_{c',d'} u(c,c';d,d';\theta) + \beta \sum_{\theta' \in \Theta} \pi_{\theta\theta'} V(c',d',\theta'),$$
$$\text{s.t. } (c,c';d,d';\theta) \in \Gamma,$$

where the domain of V is $\Omega \times \Theta$. We assume that $u(\cdot, c'; d, d'; \theta)$ and $u(c, \cdot; d, d'; \theta)$ are differentiable.

Suppose that $C(c,d,\theta)$ and $D(c,d,\theta)$ are optimal choices at the state (c,d,θ).

Definition 5. The set of feasible one-shot deviations from the optimal policies at state (c,d,θ) is

$$\Lambda(c,d,\theta) = \left\{ c' : (c,c';d,\bar{d}';\theta) \in \Gamma, \text{ and for all } \theta', \ (c', \bar{c}''(\theta'); d', \bar{d}''(\theta'); \theta') \in \Gamma \right\},$$

where $(\bar{c}', \bar{d}') = (C(c,d,\theta), D(c,d,\theta))$, and $(\bar{c}''(\theta'), \bar{d}'(\theta')) = (C(\bar{c}', \bar{d}', \theta'), D(\bar{c}', \bar{d}', \theta'))$.

[8] The bouncing ball function is the upper envelope of a set of parabolas, $\{v(\cdot, d)\}_{d \in D}$ where $v(c,d) = -\frac{2}{|d|}(c-d)(c-\frac{1}{2}d)$ and $D = \{s2^{-n} : s \in \{-1,1\}, n \in \mathbb{N}\}$. On the relevant parts of the domains, their derivatives lie in $[-1,1]$.

Corollary 3. *Let* $(\hat{c}', \hat{d}') = (C(c, d, \theta), D(c, d, \theta))$ *be optimal choices at state* (c, d, θ)*. If* $\hat{c}'$ *is an interior choice, i.e.* $\hat{c}' \in \text{int}(\Lambda(c, d, \theta))$*, then (i)* $V(\cdot, \hat{d}')$ *is differentiable at* $\hat{c}'$ *and (ii)* $\hat{c}'$ *satisfies the first-order condition*

$$-u_{c'}(c, \hat{c}'; d, \hat{d}'; \theta) = \beta \sum_{\theta'} \pi_{\theta\theta'} V_c(\hat{c}', \hat{d}', \theta') = \beta \sum_{\theta'} \pi_{\theta\theta'} u_c(\hat{c}', \hat{c}''(\theta'); \hat{d}', \hat{d}''(\theta'); \theta'),$$

where $(\hat{c}''(\theta'), \hat{d}'(\theta'))$ *are shorthand for* $(C(\hat{c}', \hat{d}', \theta'), D(\hat{c}', \hat{d}', \theta'))$*.*

Proof. We assumed that $\hat{c}'$ maximises

$$\phi(c') = u(c, c'; d, \hat{d}'; \theta) + \beta \sum_{\theta' \in \Theta} \pi_{\theta\theta'} V(c', \hat{d}', \theta'), \tag{15}$$

where the domain of ϕ is $\Lambda(c, d, \theta)$. The value function V has a differentiable lower support function at $(\hat{c}', \hat{d}', \theta')$,

$$L(c'; \hat{c}', \hat{d}', \theta') = u(c', \hat{c}''(\theta'); d', \hat{d}''(\theta'), \theta') + \beta \sum_{\theta'' \in \Theta} \pi_{\theta'\theta''} V(\hat{c}''(\theta'), \hat{d}''(\theta'), \theta''). \tag{16}$$

Therefore, ϕ is sandwiched between a corresponding differentiable lower support function, and a constant upper support function at $\hat{c}'$ By the Differentiable Sandwich Lemma, ϕ is differentiable at $\hat{c}'$. By the addition rule of reverse calculus, V is differentiable at each $(\hat{c}', \hat{d}', \theta')$. □

4.3 Policy Functions in Smooth Concave Problems

Previous theorems about the differentiability of policy functions focus on smooth deterministic concave settings. Our main contribution is that our recipe can be applied without any of these conditions. For example, in our Stackelberg illustration in Example 1, we accommodate boundaries and discontinuous marginal utilities. In optimal stopping problems (such as the unsecured credit application in Section 5.1), we accommodate discrete choices and shocks.

Araújo and Scheinkman (1977) study the differentiability of policy functions, further restricting attention to deterministic one-good problems in which the utility function is twice differentiable. Their proof is based on generalising the implicit function theorem to infinite-dimensional spaces. Thus, their approach is similar to our use of the implicit function theorem in Example 1. Reverse calculus is necessary to extend this approach beyond one endogenous function.

Santos (1991) drops the one-good restriction, but retains the concavity and twice-differentiability conditions. As mentioned above, he begins by noticing that differentiability of policy functions is equivalent to twice-differentiability of value functions. He proceeds to construct a convergent sequence of quadratic approximations of the value function to establish this twice-differentiability. Like Araújo and Scheinkman (1977), his results only address the differentiability of policy functions and value functions. He does not accommodate qualitatively different endogenous functions such as the cut-off policies in Section 5.1.

5 Applications

5.1 Unsecured Credit

Our first application is about unsecured debt contracts where borrowers may decide to either repay in full or to default. We focus on markets without collateral such as sovereign debt markets. The punishment for default is exclusion from the credit market thereafter. Nevertheless, default occasionally occurs so interest paid by the borrower must compensate for the default risk.[9] For this reason, the interest charged is non-linear and determined by a recursive relationship with the borrower's value function. If the interest rates are low, then the borrower's value of honouring debt contracts is high because rolling over debt is cheap. Conversely, if the borrower's value of repaying is high tomorrow, then the default risk today is low. This recursive relationship determines interest rates as a function of loan size and the credit limit.

The borrower's decision problem is poorly behaved for two related reasons. First, the discrete repayment choice leads to jumps in the marginal value of owing debt. Second, following marginal changes in debt, these jumps lead to kinks in the default risk and hence kinks in the interest rate. In other words, neither the value function nor the budget constraint are globally differentiable. Nevertheless, we apply our envelope theorem to establish that both endogenous functions – the value function and the interest rate – are differentiable at optimal debt choices (except for choices at the endogenous risk-free credit limit). Hence, first-order conditions apply and we can establish an Euler equation involving a marginal interest rate and a marginal continuation value. We then apply our envelope theorem to characterise the borrower's credit limit and reach our conclusion that the borrower never exhausts his endogenous credit limit.

We build on the unsecured credit analysis by Arellano (2008) which is in the tradition of Eaton and Gersovitz (1981). Arellano carefully analyses it theoretically and numerically. She also sketches a Laffer curve for the debt choice, but – without first-order conditions – does not characterise borrower behaviour along it. The following three papers apply some Euler equations, with the first explicitly acknowledging that they lack justification for differentiating the interest rates

[9] Default need not be inefficient compared to risk-free debt, as it implements risk-sharing.

with respect to loan size. We provide a justification. Aguiar and Gopinath (2006) dropped a detailed discussion of their heuristic (but now verified) Euler equation from their NBER working paper version. Similarly, we verify the heuristic Euler equations that Arellano and Ramanarayanan (2012) use to compare maturity structures of loans. Finally, Hatchondo and Martinez (2009) discuss an Euler equation, implicitly assuming differentiability of interest rates. None of these papers use first-order conditions to investigate credit limits, nor deduce our result that borrowers never exhaust their credit limits.

Model. A risk-averse borrower has a differentiable utility function u and discount factor $\beta \in (0,1)$. The borrower's marginal value of consumption at zero is infinite, i.e. $\lim_{c \to 0^+} u_1(c) = \infty$. Every period, the borrower receives an endowment x which is independently and identically distributed with density $f(\cdot)$ on the support $[x^{\min}, x^{\max}]$. We assume the borrower's endowment is bounded away from zero, i.e. $x^{\min} > 0$. To smooth out endowment shocks, the borrower may take out loans from a lender with deep pockets. We focus our attention on debt contracts of the following form. The borrower offers to pay a lender b' in the following period, but only promises to honour this debt obligation if the endowment tomorrow, x', lies in the set H'. Thus, a debt contract consists of (b', H'), both of which are chosen by the borrower. The lender is risk-neutral, discounts time at the same rate, and is therefore willing to pay $\beta \int_{H'} f(x')\, dx' b'$ in return for the promise. If the borrower defaults, he is excluded from credit markets thereafter. We also accommodate an additional exogenous sanction of $s \geq 0$ units of consumption every period for defaulting, which reflects the difficulty of settling non-financial transactions without credit.[10] The borrower's autarky value after defaulting is

$$W_{\text{aut}}(x) = u(x - s) + \beta \int_{[x^{\min}, x^{\max}]} W_{\text{aut}}(x') f(x') dx'. \tag{17}$$

The lender only agrees to the contract (b', H') if the borrower has an incentive to honour the promise for the proposed endowments H'. Specifically, the borrower's value of repaying b' at an honour endowment $x' \in H'$, denoted $W_{\text{hon}}(b', x')$, should not be less than the autarky value $W_{\text{aut}}(x)$. The borrower's value of honouring

[10] Exogenous sanctions are often included in unsecured credit models, so we include them to show the generality of our technique. Without them, Bulow and Rogoff (1989) show that exclusion from credit markets alone is an insufficient punishment for enforcing debt contracts if the borrower can make private investments.

debts is therefore[11]

$$
\begin{aligned}
W_{\text{hon}}(b,x) = \max_{c,b',H'} \; & u(c) + \int_{[x^{\min},x^{\max}]} \max\left\{W_{\text{aut}}(x'), W_{\text{hon}}(b',x')\right\} f(x') dx', \\
\text{s.t.} \quad & c + b = x + \left[\beta \int_{H'} f(x')\, dx'\right] b', \\
& W_{\text{hon}}(b',x') \geq W_{\text{aut}}(x') \text{ for all } x' \in H', \\
& b' \leq b^{\text{ponzi}}.
\end{aligned} \tag{18}
$$

The last constraint rules out Ponzi schemes and the b^{ponzi} parameter may be arbitrarily large.

Reformulation. We reformulate this problem by making two simplifications. First, Arellano (2008, Proposition 3) established that because x is IID, the honour set H' chosen by the borrower is determined by a cut-off rule $y(\cdot)$ so that the borrower honours his debt at state (b', x') if and only if $x' \geq y(b')$. In other words, the borrower only ever chooses debt contracts of the form $(b', H') = (b', [y(b'), x^{\max}])$, so debt contracts are characterised by b' alone. This means we may denote the price of debt $q(b')$ as a function of b'. Second, we substitute the budget constraint into the objective, so that the borrower's only choice is his future debt obligation b'. The reformulated problem becomes

$$
W_{\text{hon}}(b,x) = \max_{b' \leq b^{\text{ponzi}}} u(x + q(b')b' - b) + \beta W(b'), \tag{19}
$$

where

$$
W(b') = \int_{[x^{\min},x^{\max}]} \max\left\{W_{\text{aut}}(x'), W_{\text{hon}}(b',x')\right\} f(x') dx', \tag{20a}
$$

$$
q(b') = \beta[1 - F(y(b'))], \tag{20b}
$$

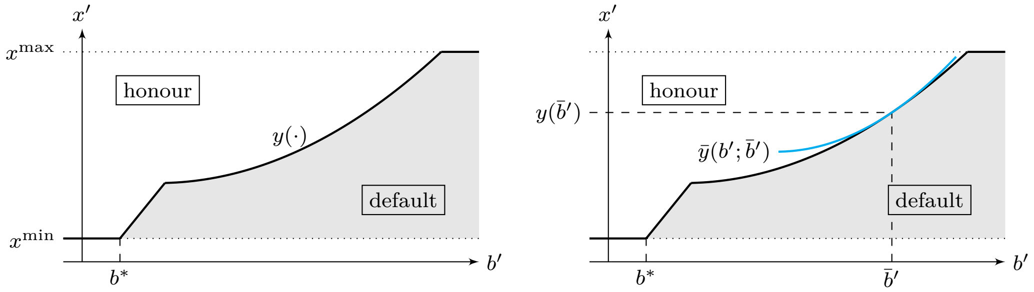
$$
y(b') = \min\left(\left\{x' \in [x^{\min}, x^{\max}] : W_{\text{hon}}(b',x') \geq W_{\text{aut}}(x')\right\} \cup \left\{x^{\max}\right\}\right). \tag{20c}
$$

We denote optimal policy functions by $\hat{b}'(b,x)$.[12]

The objective (19) has two endogenous functions, q and W, which we will show are not globally differentiable. The value function has downward kinks at states of indifference between honouring and defaulting, as in the value function of the indivisible labour choice illustration. Moreover, we have no a priori knowledge of the differentiability of the debt price. We will construct differentiable lower support functions for q and W and hence show that they both do not exhibit upward kinks at any choice, with one exception: The debt price exhibits an upward kink at the risk-free credit limit.

[11] We mention some technicalities: (i) the borrower should be constrained to choosing a measurable honour set, and (ii) the Bellman operator is well-defined for continuous value functions.

[12] The borrower might be indifferent between several optimal policies.



(a) Borrowers default when $x' < y(b')$

(b) A "lazy" borrower undervalues honouring debts, and defaults too much

Figure 8: The default cut-off rule

Differentiable Lower Support Functions. The problem of constructing a differentiable lower support function for the debt price $q(\cdot)$ is equivalent to that of constructing a differentiable upper support function for the cut-off rule $y(\cdot)$, illustrated in Figure 8a. For debts below some threshold b^*, the borrower always honours his obligations, so the cut-off $y(\cdot)$ is constant and hence differentiable on $[-\infty, b^*)$. At each debt level $\bar b' > b^*$, we now construct a differentiable upper support function for $y(\cdot)$. We consider a lazy borrower that – as a consequence of his laziness – undervalues honouring debts, and hence uses a higher cut-off than $y(\cdot)$. Specifically we consider a lazy borrower who incorrectly anticipates the state to be $(b', x') = (\bar b', y(\bar b'))$, i.e. he anticipates his state will be on the cut-off. In unanticipated states, he chooses his debt to be $\hat b''(\bar b', y(\bar b'))$ independently of the realized endowment x'. His consumption is adjusted by the differences from the anticipated endowment and debt. This lazy borrower's value function is

$$L(b', x'; \bar b') = u(x' - b' + q(\bar b'')\bar b'') + \beta W(\bar b''). \tag{21}$$

Since the lazy borrower undervalues honouring debts, his honour cut-off $\bar y(\cdot; \bar b')$ implicitly defined by

$$L(b', \bar y(b'; \bar b'); \bar b') = W_{\text{aut}}(\bar y(b'; \bar b')) \text{ for all } b' \tag{22}$$

provides an upper support function for the cut-off $y(\cdot)$ at $\bar b'$, depicted in Figure 8b. Since the lazy borrower's value function is differentiable, the implicit function theorem implies that $\bar y(\cdot; \bar b')$ is differentiable with $y_1(\bar b'; \bar b') > 1$ for all $\bar b' > b^*$.[13]

[13] Apply the implicit function theorem on the lazy borrower's value function to get

$$\bar y_1(\bar b'; \bar b') = \frac{u_1(\bar{c}'(\bar b', y(\bar b')))}{u_1(\bar{c}'(\bar b', y(\bar b'))) - u_1(x' - s)} > 1.$$

Thus far, we have established that the slope of the cut-off $y(\cdot)$ is zero approaching the risk-free limit b^* from the left, but greater than one approaching b^* from the right. Therefore, the cut-off has a downward kink at b^*, so it has no differentiable upper support function at this point. This means we have established:

Lemma 5. *At every $\bar{b}' \neq b^*$, there exists a differentiable upper support function $\bar{y}(\cdot;\bar{b}')$ for $y(\cdot)$, and hence a differentiable lower support function $\underline{q}(\cdot;\bar{b}')$ for $q(\cdot)$. Moreover, $y(\cdot)$ has an downward kink at b^* with $0 = y'(b^*-) < 1 < y'(b^*+)$.*

To construct a differentiable lower support function for W, we begin by constructing a differentiable lower support function for $W_{\text{hon}}(b', x')$. However, this time, we use a different lazy borrower's value function from the one used to construct (21). This time, the lazy borrower correctly anticipates x', but incorrectly anticipates b' to be $\bar{b}'$. He takes on a debt of $\bar{b}''(x') = \hat{b}''(\bar{b}', x')$ independently of his previous obligation of b'. His value function is

$$M(b', x'; \bar{b}') = u(x' - b' + q(\bar{b}''(x'))\bar{b}''(x')) + \beta W(\bar{b}''(x')). \tag{23}$$

This means that,

$$\underline{W}(b'; \bar{b}') = W_{\text{aut}}(x') + \int_{\bar{y}(b';\bar{b}')}^{x^{\max}} \left[M(b', x'; \bar{b}') - W_{\text{aut}}(x')\right] f(x')\, dx' \tag{24}$$

is a lower support function for W at $\bar{b}'$. We would like to establish that $\underline{W}(\cdot;\bar{b}')$ is differentiable. First, $M(\cdot, x'; \bar{b}')$ is continuously differentiable for all $(x', \bar{b}')$. Second, we note that without loss of generality, we may assume some optimal policy $\hat{b}''(\cdot,\cdot)$ is measurable, and hence the resulting lazy policy $\bar{b}''(\cdot)$ is also measurable.[14] Third, the measurability of the lazy policy implies that $M_1(b', \cdot; \bar{b}')$ is measurable for all $(b', \bar{b}')$. Moreover, it is possible to show that $M_1(b', \cdot; \bar{b}')$ is uniformly bounded for all b' in some open neighbourhood of $\bar{b}'$. Hence the Leibniz rule for differentiating under the integral sign implies that $\underline{W}(\cdot;\bar{b}')$ is differentiable at $b' = \bar{b}'$ with[15]

$$\underline{W}_1(b'; \bar{b}') = \int_{\bar{y}(b';\bar{b}')}^{x^{\max}} M_1(\bar{b}', x'; \bar{b}') f(x')\, dx'. \tag{25}$$

This means we have established:

Lemma 6. *At every $\bar{b}'$, there exists a differentiable lower support function $\underline{W}(\cdot;\bar{b}')$ for W.*

[14] See for example the Measurable Maximum Theorem in Aliprantis and Border (2006, Theorem 18.19).

[15] See for example Weizsäcker (2008, Theorem 4.6).

First-Order Conditions. We can now return to the original problem (19). If $\hat{b}'$ is an optimal debt choice at the state (b, x), then it maximises

$$\phi(b'; b, x) = u(x - b + q(b')b') + \beta W(b'). \tag{26}$$

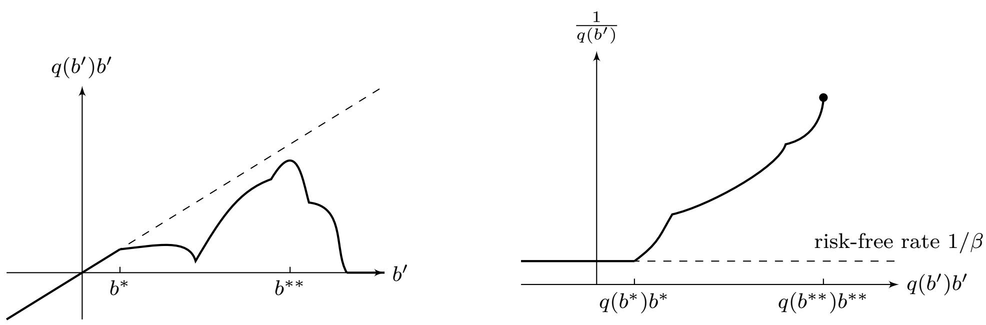
Using $\underline{q}(\cdot; \bar{b}')$ and $\underline{W}(\cdot; \bar{b}')$, we can construct a differentiable lower support for this objective at any $\bar{b}'$. By the Differentiable Sandwich Lemma (Lemma 1), the borrower's objective is differentiable at the optimal debt choice $\hat{b}'$. Moreover, by repeatedly applying the Reverse Calculus Lemma (Lemma 3), we deduce that q and W are differentiable at $\hat{b}'$. We conclude:

Corollary 4. *Suppose $\hat{b}'(\cdot, \cdot)$ is an optimal policy function, fix any state (b, x), and set $\hat{b}' = \hat{b}'(b, x)$. If $\hat{b}' \neq b^*$, then the following first-order condition holds and the endogenous functions q and V that appear in it are differentiable at $\hat{b}'$:*

$$u_1(\hat{c}(b, x))(q(\hat{b}') + q_1(\hat{b}')\hat{b}') = \beta V_1(\hat{b}') = \beta \int_{y(\hat{b}')}^{x^{max}} u_1(\hat{c}(\hat{b}', x'))f(x')\, dx', \tag{27}$$

where $\hat{c}(b, x) = x - b + q(\hat{b}'(b, x))\hat{b}'(b, x)$.

The borrower equates the marginal benefit of owing debt with the marginal cost. The marginal benefit consists of the marginal utility of consumption times the marginal revenue from promising an extra unit to the lender. The marginal cost consists of the expected marginal utility of the foregone consumption when repaying the following period (when the endowment shock is above the default cut-off).



(a) Laffer curve for debt

(b) Endogenous interest rate

Figure 9: Characterisation of endogenous borrowing

Credit Limits. We now turn our attention to the borrower's behaviour near the credit limit. The amount the lender is willing to pay, $q(b')b'$ in return for a promise of b' is not an increasing function. This is because there are two types of empty promises: $b' = 0$, and b' so large it is never honoured. The borrower's return on promises therefore follows a Laffer curve, depicted in Figure 9a. The borrower's credit limit is the maximum of this curve, $q(b^{**})b^{**}$, where

$$b^{**} = \arg\max_{b'} q(b')b'. \tag{28}$$

If $b^{**} > b^*$, then we have already constructed a differentiable lower support function for q, so the Differentiable Sandwich Lemma (Lemma 1) together with the Reverse Calculus Lemma (Lemma 3) imply that q is differentiable at b^{**} with

$$q(b^{**}) + q_1(b^{**})b^{**} = 0. \tag{29}$$

Substituting this into the Euler equation (27), we see that the marginal benefit of taking on debt at b^{**} is zero, while the marginal cost is positive. Therefore, we conclude

Corollary 5. *For any given model primitives, either*

(i) the overall and risk-free credit limits coincide, i.e. $b^{**} = b^*$, *or*

(ii) the overall credit limit is higher and exhausting it is suboptimal, i.e. $b^{**} > b^*$ *and* $\hat{b}'(b, x) < b^{**}$ *for all states* (b, x).

This conclusion is a logical generalisation of behaviour in Aiyagari's (1994) model. Both here and there, the borrower reaches the risk-free credit limit with positive probability. In the model we study, the overall credit limit is potentially higher, as the borrower has the additional possibility of taking out risky loans. However, behaviour near the two credit limits is strikingly different. Below the risk-free limit, the interest rate $1/q(b')$ remains constant as the loan size $q(b')b'$ increases. Above the risk-free limit, the interest rate increases as the borrower takes on more debt and increases the default risk, as depicted in Figure 9b. This difference accounts for why borrowers might exhaust their risk-free limit, but not their overall limit.

Arellano (2008, Figure 2) plots a similar Laffer curve as in Figure 9a. Possibly for computational reasons, her curve is smooth and does not depict the upward kink of the Laffer curve at the risk-free limit, b^*. She does not apply first-order conditions along the Laffer curve.

Final Remarks. Despite our results regarding first-order conditions, credit limits, and the Laffer curve, some questions remain. First, we do not know if the Laffer curve is single-peaked. Second, the IID shock assumption was important for Arellano (2008) to establish that the default policy is a cut-off rule. More generally, persistent shocks cause interest rates to depend on the shock in addition to the size of the loan, which is crucial for understanding how credit markets operate when borrowers are distressed. Nevertheless, we believe our analysis can be generalised. Chatterjee, Corbae, Nakajima and Ríos-Rull (2007, Theorem 3) established that persistent shocks lead to two-sided cut-off rules. We conjecture that it is possible to construct differentiable support functions for the two cut-offs, and use this to construct a differentiable upper support function for the repayment probability. Finally, we believe that first-order conditions will be central to exploring extensions of the model to study issues such as partial default and optimal term structure.

5.2 Adjustment Costs

Firms are slow to adjust prices, labour forces, and capital stocks in reaction to changes in market conditions. One explanation for this is that firms face adjustment costs such as fixed costs or other non-convex costs. There is a large literature investigating how shocks propagate in the presence of adjustment costs and whether or not adjustment costs amplify shocks; see the surveys by Khan and Thomas (2008a), Leahy (2008), and Caplin and Leahy (2010). However, most of this literature is purely empirical, because the theory of adjustment costs faces two important obstacles. One is the complexity of optimal policy functions. Both theoretical and empirical analysis has only been tractable thus far when optimal polices involve smooth cut-off rules for determining when adjustments take place.[16] The other is the difficulty in deriving recursive first-order conditions, as the value of adjustment is not differentiable in general. Caballero and Engel (1999) use shocks that enter linearly into the production function to smooth out the kinks in the value function. Under this specific structure, they are able to take first-conditions to characterise optimal adjustments. To make this operational, they conjecture that adjustments follow a smooth two-sided (S, s) policy, but only verify this numerically.[17] Gertler and Leahy (2008) study a quadratic approximation of the firm's objective function in which the non-differentiable terms in the continuation value of adjustment vanish and optimal policies are smooth two-sided (S, s). They establish low error bounds for this approximation for an appropriate range of adjustment cost and shock parameters. Elsby and Michaels (2014) use first-order conditions under the conjecture that the optimal adjustment policy is a smooth two-sided (S, s) policy, also without providing sufficient conditions on primitives for this conjecture to hold. For the purposes of illustration, Cooper and Halti-

[16] Specifically, we say that a policy is a smooth two-sided (S, s) policy if (i) for every capital (or labour or price) level, the set of shocks for which the firm makes an adjustment is an interval and (ii) the upper and lower end points of this interval are differentiable functions of the capital level.

[17] Caballero and Engel (1999, Footnote 16)

wanger (2006, Section 3.2) and Khan and Thomas (2008b, Appendix B) provide derivatives only in the absence of fixed costs; we show these derivatives hold generally. An alternative approach is to assume that information arrives gradually over continuous time; see Harrison *et al.* (1983), Stokey (2008), and Golosov and Lucas (2007).

The fundamental problem is that if a firm invests more today, then it might defer subsequent investment longer. Thus a small change in today's choice may lead to a lumpy change in a later choice, giving a non-differentiable and non-concave value of investment. We show that at optimal adjustment choices, the value function is differentiable so that recursive first-order conditions are applicable. We require only very weak assumptions on the primitives. In particular, our result remains true even when optimal policies are not two-sided (S, s) (see for example Bar-Ilan, 1990).

Model. In a general formulation, a firm is endowed with a capital stock k and shock z. Shocks evolve according to a Markov process with conditional distribution $P(z'|z)$. In each period, the firm's flow profit is $\pi(k,z)$; for example $\pi(k,z) = pf(k,z) - rk$ where p is output price, f is the production function, and r is the rental rate of capital. The firm pays an adjustment cost $c(k',k,z)$; non-adjustment is costless. We assume the flow profit $\pi(\cdot,z)$ is differentiable for all z, and that the adjustment cost $c(\cdot,\cdot,z)$ function is differentiable at all points (k,k',z) such that $k' \neq k$. For example, this accommodates the pure fixed-cost function, $c(k',k,z) = I(k' \neq k)$. The firm's value before adjusting its capital stock at state (k,z) is $V(k,z)$. Its value after adjusting its capital stock to k' is $W(k',z)$. These two value functions are related by the following two Bellman equations:

$$V(k,z) = \max_{k'} \pi(k,z) - c(k',k,z) + \beta W(k',z), \tag{30a}$$

$$W(k',z) = \int V(k',z')\, dP(z'|z). \tag{30b}$$

Our goal is to establish the first-order condition for the capital choice k'

$$c_1(k',k,z) = \beta W_1(k',z) \tag{31}$$

and to derive a formula for the marginal value of investment $W_1(k',z)$ at the optimal choice $k' = \hat{k}'(k,z)$. If there is a fixed cost of an adjustment, then this formula will only be satisfied when the agent makes an adjustment, i.e. at shocks z lying in the optimal adjustment set

$$\hat{A}(k) = \left\{ z : \hat{k}'(k,z) \neq k \right\}. \tag{32}$$

Differentiable Lower Support Functions. We construct a differentiable lower support function for the value function V by considering a lazy manager who knows the optimal policy when he begins with a familiar capital stock of $k = \bar{k}$. The obvious lazy manager policy of sticking to the same capital choice when $k \neq \bar{k}$ is not useful here, because it leads to a discontinuous lazy value function.[18] Instead, we consider a lazy manager who uses the familiar adjustment set and adjustment level for unfamiliar capital stocks, i.e. he waits until he draws a shock $z \in \hat{A}(\bar{k})$ and adjusts to $\hat{k}'(\bar{k}, z)$. Thereafter, his choices coincide with the rational manager. His value function is

$$L(k, z; \bar{k}) = \pi(k, z) + \begin{cases} \beta \int L(k, z'; \bar{k})\, dP(z'|z) & \text{if } z \notin \hat{A}(\bar{k}), \\ -c(\hat{k}'(\bar{k}, z), k, z) + \beta W(\hat{k}'(\bar{k}, z), z) & \text{if } z \in \hat{A}(\bar{k}). \end{cases} \tag{33}$$

It is straightforward to calculate the lazy manager's marginal value of capital, because the capital stock k does not affect any subsequent choices:[19]

$$L_1(k, z; \bar{k}) = \pi_1(k, z) + \begin{cases} \beta \int L_1(k, z'; \bar{k})\, dP(z'|z) & \text{if } z \notin \hat{A}(\bar{k}), \\ -c_2(\hat{k}'(\bar{k}, z), k, z) & \text{if } z \in \hat{A}(\bar{k}). \end{cases} \tag{34}$$

First-Order Conditions. If $\hat{k}'$ is an optimal choice at the state (k, z), then $\hat{k}'$ maximises

$$\phi(k'; k, z) = \pi(k, z) - c(k', k, z) + \beta W(k', z). \tag{35}$$

By substituting in (33) and (30b), we may construct a differentiable lower support function for $\phi(\cdot; k, z)$ at $\hat{k}'$. Lemma 2 provides a differentiable upper support function, so Lemma 1 establishes the following corollary.

Corollary 6. *If making an adjustment is optimal at state (k, z), i.e. $z \in \hat{A}(k)$, then the investment value W is differentiable in capital at $(\hat{k}'(k, z), z)$ and*

$$c_1(\hat{k}'(k, z), k, z) = \beta W_1(\hat{k}'(k, z), z) = \beta \int \tilde{L}_1(\hat{k}'(k, z'), z')\, dP(z'|z), \tag{36a}$$

$$\textit{where } \tilde{L}_1(k, z) = \pi_1(k, z) + \begin{cases} \beta \int \tilde{L}_1(k, z')\, dP(z'|z) & \textit{if } z \notin \hat{A}(k), \\ -c_2(\hat{k}'(k, z), k, z) & \textit{if } z \in \hat{A}(k). \end{cases} \tag{36b}$$

[18] This obvious lazy manager makes an extra adjustment even if the capital stock is only slightly different from the familiar level.

[19] The lazy manager's marginal value follows from the chain rule applied to (i) the expected discounted profit as a function of all state-contingent capital choices, holding adjustment times fixed, and (ii) the lazy capital choices as a function of initial capital k only.

The first equation says that the marginal adjustment cost should equal the marginal value of investment, which is the same for both rational and lazy managers. The second equation says that the marginal value of increasing investment equals the expected marginal increase in profit until the next adjustment plus the marginal decrease in the subsequent adjustment cost. We have thus shown that first-order conditions are generally valid even if the optimal adjustment policies are not (S,s). In other words, we have established that the applicability of first-order conditions is not an obstacle to the theoretical analysis of the implications of adjustment costs to prices, labour forces, and capital stocks. The only remaining obstacle is understanding when optimal policies are (S,s).

5.3 Social Insurance

Governments run public health, unemployment and disability insurance programs, and private companies offer insurance contracts. These are constrained by frictions such as hidden information, adverse selection, and moral hazard. Informal insurance arises within well-connected families and communities when they can partially overcome these frictions. There is a large literature studying informal insurance, and the interaction of informal insurance with other forms of insurance.[20] In the dynamic insurance models of Thomas and Worrall (1988, 1990) and Kocherlakota (1996), the main issue is how cross-subsidisation may be self-enforcing. Agents with good luck subsidise those with bad luck in return for promises of future payments and insurance. These papers study smooth convex environments in which the Benveniste and Scheinkman (1979) theorem provides a formula for the marginal cost of making promises.[21] However, some important insurance problems involve non-smooth settings. We focus on a setting similar to that of Morten (2013), which is an extension of Ligon *et al.*'s (2002) model of self-enforcing dynamic insurance. Villagers share risk among themselves by both sharing divisible output and sending some members of the community to find temporary work in cities. The temporary migration decisions are inherently discrete as they involve a fixed cost of moving to the city and back. Other examples of indivisible items in village economies include livestock, medical treatments, agricultural land (due to high legal costs), and houses. This environment is non-smooth and non-concave, so the marginal cost of promises does not exist globally. Nevertheless, our envelope theorem applies and allows us to characterise optimal insurance policies in terms of the marginal cost of

[20] Apart from the papers we discuss, Townsend (1994), Attanasio and Ríos-Rull (2000), and Krueger and Perri (2006) are important papers.

[21] Kocherlakota (1996) mistakenly claims his value function is differentiable. Koeppl (2006) amends his Bellman equation along the lines of Thomas and Worrall (1988). See also Ljungqvist and Sargent (2012, Chapter 20), and Rincón-Zapatero and Santos (2009, Section 4.2) for further discussion.

promises. Optimal policies involve sharing risk through allocating indivisible temporary work obligations; divisible consumption is then allocated to smooth out the marginal utility of consumption across states.

Model. Consider the following dynamic risk-sharing game between two households $h \in \{1, 2\}$. Each period begins with a Markov shock $s \in S$ with transition function $p(s'|s)$. The shock determines each household's endowment of a divisible consumption good, $C_h(s)$. The aggregate endowment is $C(s) = C_1(s) + C_2(s)$. In addition, each household may produce M units of the consumption good from temporary migrant work in a city. We write $d_h = 1$ if the household migrates, and $d_h = 0$ otherwise. We assume that the utility from consumption $u(\cdot, d_h)$ is differentiable, and that the marginal utility approaches infinity as consumption approaches zero. The autarky value of each household is

$$V_h^{\text{aut}}(s) = \max_{d_h} u(C_h(s) + Md_h, d_h) + \beta \sum_{s'} p(s'|s) V_h^{\text{aut}}(s'). \tag{37}$$

Before investigating the social insurance arrangements with autarky constraints, we present the social planner's problem with Negishi weights η_1 and η_2:

$$W(s) = \max_{c_1, d_1} \eta_1 u(c_1, d_1) + \eta_2 u(c_2, d_2) + \beta \sum_{s' \in S} p(s'|s) W(s') \tag{38a}$$

$$\text{where} \quad c_1(s) + c_2(s) = C(S) + (d_1 + d_2)M. \tag{38b}$$

The first-order condition with respect to c_1 gives the Borch (1962) equation

$$\frac{u_1(c_1, d_1)}{u_1(c_2, d_2)} = \frac{\eta_2}{\eta_1}. \tag{39}$$

This means that after the social planner allocates the migration decisions, she adjusts the consumption good until the planner's marginal rate of substitution between the households is equal to the ratio of Negishi weights at all states and dates.

Now, we add in autarky constraints to study the optimal incentive-compatible social insurance contract. The value function for household 1 can be formulated recursively in terms of a principal-agent problem in which household 1 acts as an insurer and is able to promise future utility to household 2. This promised utility is a state variable, and has a corresponding promise-keeping constraint. Both households can leave the contract at any time, so there is an autarky constraint for each of them.

$$
\begin{aligned}
V(s,v_2) = \max_{c_1,d_1,d_2,v_2'(s')} \ & u(c_1,d_1) + \beta \sum_{s'\in S} p(s'|s) V(s', v_2'(s')) \\
\text{s.t.} \quad & (\mathrm{PK}_2) \quad u(c_2,d_2) + \beta \sum_{s'\in S} p(s'|s) v_2'(s') = v_2, \\
& (\mathrm{A}_1\text{-}s') \quad V(s', v_2'(s')) \geq V_1^{\mathrm{aut}}(s') \text{ for all } s' \in S, \\
& (\mathrm{A}_2\text{-}s') \quad v_2'(s') \geq V_2^{\mathrm{aut}}(s') \text{ for all } s' \in S, \\
\text{where} \quad & c_2 = C(s) + (d_1 + d_2)M - c_1.
\end{aligned} \tag{40}
$$

First-Order Conditions. We now derive first-order conditions by applying Corollary 3. For the purposes of applying the corollary, the discrete choice is the migration allocation (d_1, d_2), and the continuous choices are the state-contingent promised utilities $v_2'(\cdot)$. The consumption choice c_1 can be eliminated by substituting in the promise-keeping constraint. Corollary 3 delivers the following:

Corollary 7. *Let $(\hat{d}_1(s,v_2), \hat{d}_2(s,v_2), \hat{v}_2'(s'|s,v_2))$ be an optimal choice at state (s,v_2). Fix some state (s,v_2) and some future shock s'. If neither autarky constraint binds for the choice of $\hat{v}_2'(s'|s,v_2)$, then the value function $V(s',\cdot)$ is differentiable at $\hat{v}_2'(s'|s,v_2)$ with*

$$
-\frac{u_1(\hat{c}_1(s,v_2), \hat{d}_1(s,v_2))}{u_1(\hat{c}_2(s,v_2), \hat{d}_2(s,v_2))} = V_2(s', \hat{v}_2'(s'|s,v_2)) \tag{41}
$$

$$
= -\frac{u_1(\hat{c}_1(s', \hat{v}_2'(s'|s,v_2)), \hat{d}_1'(s', \hat{v}_2'(s'|s,v_2)))}{u_1(\hat{c}_2(s', \hat{v}_2'(s'|s,v_2)), \hat{d}_2'(s', \hat{v}_2'(s'|s,v_2)))}, \tag{42}
$$

where $\hat{c}_1(s,v_2)$ and $\hat{c}_2(s,v_2)$ are implicitly defined in terms of the other optimal choices via the promising-keeping constraints.

This equation is the Borch (1962) equation which characterises perfect insurance – the social planner's marginal rate of substitution is equated across states and time periods. This means we have shown that with both divisible and indivisible choices, there is perfect insurance between households at all states and times for which the autarky constraints are lax. When an autarky constraint binds, the Negishi weights are adjusted and perfect insurance continues until an autarky constraint binds in the future. This generalises the conclusion drawn by Thomas and Worrall (1988) when indivisible choices are absent.

6 Conclusion

All envelope theorems have a sandwich idea at their core. Previous proofs were structured around sandwiches of inequalities of directional derivatives. By restructuring around sandwiches of differentiable upper and lower support functions, we gain two things. First, we do not require any of the strong technical conditions from previous envelope theorems, and can accommodate primitives with Inada conditions. Second and more importantly, our approach potentially applies to any type of endogenous functions that might need to be differentiated in a first-order condition.

Our method gains us a straightforward way of mixing and matching different constructions of upper and lower halves of sandwiches. We used five constructions throughout, namely (i) horizontal lines above maxima, (ii) supporting hyperplanes above concave functions, (iii) reverse calculus, (iv) lazy value functions below rational value functions, and (v) lazy cut-off rules. Of these, only the reverse calculus construction is truly unprecedented. The power of our approach derives from the ability to combine these constructions. For example, the unsecured credit application uses all but the supporting hyperplane construction. There are also other possibilities that we did not explore. Decision makers can be "lazy" in ways that lead to upper support functions, such as being lazily optimistic about future opportunities. In bargaining games, a lower support function for one player's value function leads to an upper support function for the other player's value function.

To conclude, our new approach reveals that trade-offs which previously seemed poorly behaved in fact have smooth structures within them that lead to first-order characterisations of optimal decisions.

A Support Functions and Subdifferentials

The notion of a differentiable lower support function generalises the classic ideas from convex analysis of supporting hyperplanes and subdifferentials. In this appendix, we establish a tight equivalence between differentiable lower support functions and Fréchet subdifferentials. These were once seen as a promising way to generalise the classical techniques of convex optimisation described by Rockafellar (1970) beyond convex settings. However, according to Kruger (2003), these were abandoned because of "rather poor calculus" as Fréchet subdifferentials do not sum, i.e. $\partial_F(f+g)(x) \neq \partial_F f(x) + \partial_F g(x)$. In light of our developments, we believe that Fréchet subdifferentials may have other applications to optimisation theory.

Suppose $(X, \|\cdot\|)$ is a Banach space and $C \subseteq X$.

Definition 6. A function $f : C \to \mathbb{R}$ is **Fréchet subdifferentiable** at $\bar{c}$ if there is some $m^* \in C^*$ such that

$$\liminf_{\Delta c \to 0} \frac{f(\bar{c} + \Delta c) - f(\bar{c}) - m^* \Delta c}{\|\Delta c\|} \geq 0. \tag{43}$$

Such an m^* is called a Fréchet subderivative of f at $\bar{c}$, and the set of all subderivatives is called the Fréchet subdifferential of f at $\bar{c}$, denoted $\partial_F f(\bar{c})$. Definitions for Fréchet superdifferentiable, superderivatives, and superdifferentials are analogous.

Theorem 2. *m^* is a Fréchet subderivative of $f : C \to \mathbb{R}$ at $\bar{c}$ if and only if f has a differentiable lower support function L at $\bar{c}$ such that $L_1(\bar{c}) = m^*$.*

Proof. If L is such a differentiable lower support function, then $L_1(\bar{c}) = m^*$, i.e.

$$\lim_{\Delta c \to 0} \frac{L(\bar{c} + \Delta c) - f(\bar{c}) - m^* \Delta c}{\|\Delta c\|} = 0. \tag{44}$$

Since $f(\bar{c} + \Delta c) \geq L(\bar{c} + \Delta c)$ for all Δc, it follows that

$$\liminf_{\Delta c \to 0} \frac{f(\bar{c} + \Delta c) - f(\bar{c}) - m^* \Delta c}{\|\Delta c\|} \geq 0 \tag{45}$$

and hence m^* is a Fréchet subderivative of f at $\bar{c}$.

Conversely, suppose that m^* is a subderivative of f at $\bar{c}$. We claim that

$$L(c) = \min \{f(c), f(\bar{c}) + m^*(c - \bar{c})\} \tag{46}$$

is a differentiable lower support function of f at $\bar{c}$. By construction, L is a lower support function. Moreover, the function $U(c) = f(\bar{c}) + m^*(c - \bar{c})$ is a differentiable upper support function of L at $\bar{c}$; by the first part of the theorem, $U_1(\bar{c}) = m^*$ is a superderivative of L at $\bar{c}$. On the other side, m^* is a subderivative of L at $\bar{c}$ because

$$\begin{aligned}
&\liminf_{\Delta c \to 0} \frac{L(\bar{c} + \Delta c) - f(\bar{c}) - m^* \Delta c}{\|\Delta c\|} \\
&= \min \left\{ 0, \liminf_{\Delta c \to 0} \frac{f(\bar{c} + \Delta c) - f(\bar{c}) - m^* \Delta c}{\|\Delta c\|} \right\} \\
&\geq 0.
\end{aligned} \tag{47}$$

Therefore, L is differentiable at $\bar{c}$ with $L_1(\bar{c}) = m^*$. □

Lemma 1 then becomes a classic result.

Lemma 7. *If m^* is a Fréchet subderivative of $f : C \to \mathbb{R}$ at $\bar{c}$ and M^* is a superderivative of f at $\bar{c}$, then f is differentiable at $\bar{c}$ with $f'(\bar{c}) = m^* = M^*$.*

B Further Reverse Calculus Rules

This appendix provides two further reverse calculus rules that were not used in the paper, but might be useful for other problems. Specifically, the rules relate to convex combinations and function composition.

The rule for convex combinations is complicated, because the forward calculus step is not obvious. The following lemma incorporates both a forward and reverse calculus result.

Lemma 8. *Suppose $E : C \to [0,1]$, $F : C \to \mathbb{R}$, and $G : C \to \mathbb{R}$ have differentiable lower support functions e, f, and g respectively at $\bar{c}$. Consider the function*

$$H(c) = E(c)F(c) + (1 - E(c))G(c).$$

If $F(\bar{c}) > G(\bar{c})$, then

(i) The function $h(c) = e(c)f(c) + (1 - e(c))g(c)$ is a differentiable (local) lower support function for H at $\bar{c}$.

(ii) If H is differentiable at $\bar{c}$, then e, f, and g are also differentiable at $\bar{c}$.

Proof. Consider the two functions,

$$\begin{aligned} h(c) &= e(c)f(c) + (1 - e(c))g(c) \\ \tilde{h}(c) &= E(c)f(c) + (1 - E(c))g(c). \end{aligned}$$

Since $f(\bar{c}) > g(\bar{c})$, we have that $h(c) \leq \tilde{h}(c)$, and hence $h(c) \leq H(c)$ in some open neighbourhood of $\bar{c}$. This establishes part (i).

For part (ii), we see that $\tilde{h}$ is differentiably sandwiched between h and H at $\bar{c}$. By the Differentiable Sandwich Lemma, $\tilde{h}$ are differentiable at $\bar{c}$. This implies $E(c) = [\tilde{h}(c) - g(c)]/[f(c) - g(c)]$ is also differentiable at $\bar{c}$. Therefore, both terms of H, namely $E(c)F(c)$ and $(1 - E(c))G(c)$, have differentiable lower support functions, $E(c)f(c)$ and $(1 - E(c))g(c)$, respectively. Part (i) of Lemma 3 implies that both terms are differentiable at $\bar{c}$, and hence F and G are differentiable at $\bar{c}$. □

Finally, we consider function composition of two endogenous functions.

Lemma 9. *If $H(c) = J(K(c))$ is differentiable at $\bar{c}$, where*

- *$J : \mathbb{R} \to \mathbb{R}$ has an inverse J^{-1} and a differentiable lower support function $j(\cdot)$ at $K(\bar{c})$,*
- *$K : \mathbb{R} \to \mathbb{R}$ has an inverse K^{-1} and a differentiable lower support function $k(\cdot)$ at $\bar{c}$, and*

- $j'(K(\bar{c})) \neq 0$ *and* $k'(\bar{c}) \neq 0$,

then J *and* K *are differentiable at* $K(\bar{c})$ *and* $\bar{c}$ *respectively.*

Proof. We assume without loss of generality that $j'(K(\bar{c})) > 0$.[22] We now establish that this implies j^{-1} is a differentiable *upper* support function for J^{-1}. To see this, we evaluate the inequality $j(c) \leq J(c)$ at $J^{-1}(x)$ which gives

$$j(J^{-1}(x)) \leq J(J^{-1}(x)) = x.$$

Applying j^{-1} to both sides gives $J^{-1}(x) \leq j^{-1}(x)$.

We can express $K(\cdot)$ as a function of J and H as follows:

$$J^{-1}(H(c)) = J^{-1}(J(K(c))) = K(c).$$

This has a differentiable upper support function $j^{-1}(H(c))$ at $\bar{c}$. Thus K has differentiable upper and lower support functions at $\bar{c}$, and is therefore differentiable by Lemma 1. Next, evaluating $H(c) = J(K(c))$ at $c = K^{-1}(x)$ gives

$$H(K^{-1}(x)) = J(K(K^{-1}(x))) = J(x),$$

so J is differentiable at $K(\bar{c})$ by the chain rule and inverse function theorem. □

References

Aguiar, M. and Gopinath, G. (2006). Defaultable debt, interest rates and the current account. *Journal of International Economics*, **69** (1), 64–83.

Aiyagari, S. R. (1994). Uninsured idiosyncratic risk and aggregate saving. *Quarterly Journal of Economics*, **109** (3), 659–684.

Aliprantis, C. D. and Border, K. C. (2006). *Infinite Dimensional Analysis*. Springer Verlag, 3rd edn.

Amir, R., Mirman, L. J. and Perkins, W. R. (1991). One-sector nonclassical optimal growth: Optimality conditions and comparative dynamics. *International Economic Review*, **32** (3), 625–644.

Araújo, A. and Scheinkman, J. A. (1977). Smoothness, comparative dynamics, and the turnpike property. *Econometrica*, **45** (3), 601–620.

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

[22] If $j'(K(\bar{c})) < 0$, then the lemma can be applied to $\tilde{H}(c) = \tilde{J}(K(c))$, where $\tilde{H}(c) = -\frac{1}{H(c)+x_0}$ and $\tilde{J}(c) = -\frac{1}{J(c)+x_0}$, where x_0 is a suitable constant to prevent division by zero near $K(\bar{c})$. In this case, $\tilde{j}(c) = -\frac{1}{j(c)+x_0}$ is a lower support function for $\tilde{J}$ with a strictly positive derivative $\tilde{j}'(c) = \frac{1}{[j(c)+x_0]^2} j'(c)$.

— and RAMANARAYANAN, A. (2012). Default and the maturity structure in sovereign bonds. *Journal of Political Economy*, **120** (2), 187–232.

ATTANASIO, O. and RÍOS-RULL, J.-V. (2000). Consumption smoothing in island economies: Can public insurance reduce welfare? *European Economic Review*, **44** (7), 1225–1258.

BAR-ILAN, A. (1990). Trigger-target rules need not be optimal with fixed adjustment costs: A simple comment on optimal money holding under uncertainty. *International Economic Review*, **31** (1), 229–234.

BENVENISTE, L. M. and SCHEINKMAN, J. A. (1979). On the differentiability of the value function in dynamic models of economics. *Econometrica*, **47** (3), 727–732.

BORCH, K. (1962). Equilibrium in a reinsurance market. *Econometrica*, **30** (3), 424–444.

BULOW, J. and ROGOFF, K. (1989). Sovereign debt: Is to forgive to forget? *American Economic Review*, **79** (1), 43–50.

CABALLERO, R. J. and ENGEL, E. M. R. A. (1999). Explaining investment dynamics in U.S. manufacturing: A generalized (S,s) approach. *Econometrica*, **67** (4), 783–826.

CAPLIN, A. and LEAHY, J. (2010). Economic theory and the world of practice: A celebration of the (S, s) model. *Journal of Economic Perspectives*, **24** (1), 183–201.

CHATTERJEE, S., CORBAE, D., NAKAJIMA, M. and RÍOS-RULL, J.-V. (2007). A quantitative theory of unsecured consumer credit with risk of default. *Econometrica*, **75** (6), 1525–1589.

CLARKE, F. H. (1975). Generalized gradients and applications. *Transactions of The American Mathematical Society*, **205**, 247–262.

COOPER, R. W. and HALTIWANGER, J. C. (2006). On the nature of capital adjustment cost. *Review of Economic Studies*, **73**, 611–633.

EATON, J. and GERSOVITZ, M. (1981). Debt with potential repudiation: Theoretical and empirical analysis. *Review of Economic Studies*, **48** (2), 289–309.

ELSBY, M. and MICHAELS, R. (2014). Fixed adjustment costs and aggregate fluctuations, mimeo.

GERTLER, M. and LEAHY, J. (2008). A Phillips curve with an Ss foundation. *Journal of Political Economy*, **116** (3), 533–572.

GOLOSOV, M. and LUCAS, R. E. (2007). Menu costs and Phillips curves. *Journal of Political Economy*, **115** (2), 171–199.

HARRISON, J. M., SELLKE, T. M. and TAYLOR, A. J. (1983). Impulse control of brownian motion. *Mathematics of Operations Research*, **8**, 454–466.

HATCHONDO, J. C. and MARTINEZ, L. (2009). Long-duration bonds and sovereign defaults. *Journal of International Economics*, **79** (1), 117–125.

KHAN, A. and THOMAS, J. K. (2008a). Adjustment costs. In S. N. Durlauf and L. E. Blume (eds.), *The New Palgrave Dictionary of Economics*, Basingstoke: Palgrave Macmillan.

— and — (2008b). Idiosyncratic shocks and the role of nonconvexities in plant and aggregate investment dynamics. *Econometrica*, **76** (2), 395–436.

KOCHERLAKOTA, N. R. (1996). Implications of efficient risk sharing without commitment. *Review of Economic Studies*, **63** (4), 595–609.

KOEPPL, T. V. (2006). Differentiability of the efficient frontier when commitment to risk sharing is limited. *B.E. Journal of Macroeconomics*, **6** (1), 1–6.

KRUEGER, D. and PERRI, F. (2006). Does income inequality lead to consumption inequality: Evidence and theory. *Review of Economic Studies*, **76**, 163–193.

KRUGER, A. Y. (2003). On Fréchet subdifferentials. *Journal of Mathematical Sciences*, **116**, 3325–3358.

KYDLAND, F. E. and PRESCOTT, E. C. (1977). Rules rather than discretion: The inconsistency of optimal plans. *Journal of Political Economy*, **85** (3), 473–491.

LEAHY, J. (2008). s-S models. In S. N. Durlauf and L. E. Blume (eds.), *The New Palgrave Dictionary of Economics*, Basingstoke: Palgrave Macmillan.

LIGON, E., THOMAS, J. P. and WORRALL, T. (2002). Informal insurance arrangements with limited commitment: Theory and evidence from village economies. *Review of Economic Studies*, **69** (1), 209–244.

LJUNGQVIST, L. and SARGENT, T. J. (2012). *Recursive Macroeconomic Theory.* MIT Press, 3rd edn.

MILGROM, P. and SEGAL, I. (2002). Envelope theorems for arbitrary choice sets. *Econometrica*, **70** (2), 583–601.

MIRMAN, L. J. and ZILCHA, I. (1975). On optimal growth under uncertainty. *Journal of Economic Theory*, **11** (3), 329–339.

MORTEN, M. (2013). Temporary migration and endogenous risk sharing in village India, mimeo.

RINCÓN-ZAPATERO, J. P. and SANTOS, M. S. (2009). Differentiability of the value function without interiority assumptions. *Journal of Economic Theory*, **144** (5), 1948–1964.

ROCKAFELLAR, R. T. (1970). *Convex Analysis.* Princeton University Press.

SANTOS, M. S. (1991). Smoothness of the policy function in discrete time economic models. *Econometrica*, **59** (5), 1365–1382.

STOKEY, N. L. (2008). *The Economics of Inaction: Stochastic Control models with fixed costs.* Princeton University Press.

— and LUCAS, R. E. (1989). *Recursive Methods in Economic Dynamics.* Harvard University Press.

THOMAS, J. and WORRALL, T. (1988). Self-enforcing wage contracts. *Review of Economic Studies*, **55** (4), 541–553.

— and — (1990). Income fluctuation and asymmetric information: An example of a repeated principal-agent problem. *Journal of Economic Theory*, **51** (2), 367–390.

TOWNSEND, R. M. (1994). Risk and insurance in village India. *Econometrica*, **62** (3), 539–591.

VARIAN, H. R. (1992). *Microeconomic Analysis.* Norton, 3rd edn.

WEIZSÄCKER, H. V. (2008). Basic measure theory, lecture notes.