

Reverse Calculus and Nested Optimization^{*}

Andrew Clausen
University of Edinburgh

Carlo Strub
University of St Gallen

23rd April 2019

Abstract

Nested optimization problems arise when an agent must take into account the effect of their decisions on their own future behavior, or the behavior of others. In these problems, calculating marginal costs and benefits involves differentiating the solutions to nested problems. But are these solutions differentiable functions? We develop a tool called *Reverse Calculus*, and establish first-order conditions for (i) a Stackelberg leader considering the follower’s best response function, (ii) a sovereign borrower considering its own future default policy, and (iii) non-convex dynamic programming problems.

Keywords: first-order conditions; non-convex dynamic programming; Stackelberg problems; unsecured credit

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, the marginal benefit of preparing for the future is endogenous, and envelope theorems have established its existence in well-behaved convex settings. But there are

^{*}Earlier versions of this paper circulated under the titles “Envelope Theorems for Non-Smooth and Non-Concave Optimization” and “A General and Intuitive Envelope Theorem.” 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, José-Victor Ríos-Rull, Felipe Saffie, József Sákovics, Ilya Segal, Shouyong Shi, Philip Short, Jonathan Thomas, Xi Weng, Tim Worrall, and Mark Wright for fruitful discussions. We especially thank an anonymous referee for suggesting we restructure our paper around the Differentiable Sandwich Lemma. Email: andrew.clausen@ed.ac.uk and cs@carlostrub.ch.

many other important economic problems in which policy or value functions appear in optimization objectives or constraints. In these nested optimization problems, it is unknown whether first-order conditions are applicable.

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 3.1](#). 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 involving marginal costs and marginal benefits.

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 3.2](#). First, the borrower's future default policy appears in his objective, because it determines default risk and hence interest rates. 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 existing tools.

Techniques. We devise a recipe that addresses these threats to first-order conditions. The recipe makes use of two new techniques. Our **Differentiable Sandwich Lemma** reformulates a classical result about subderivatives in simple form that is useful for establishing first-order conditions.¹ 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 1a. 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 accommodates functions that are neither continuous nor differentiable in any open neighbourhood of \bar{c} , as depicted in Figure 1b.

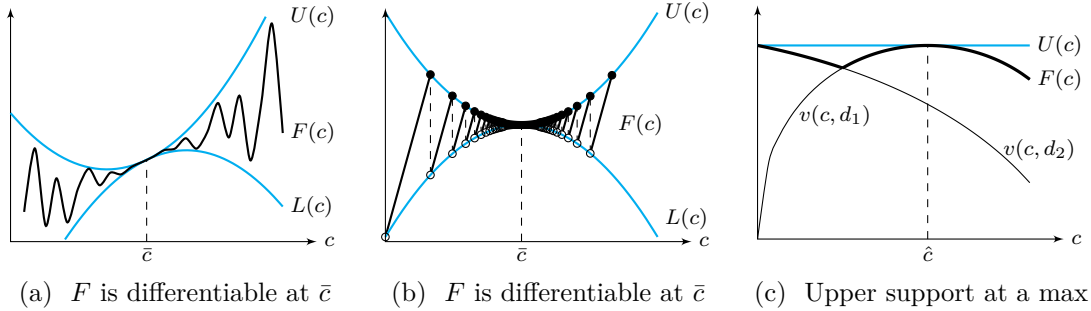


Figure 1: Differentiable Sandwich Lemma

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 $v(c, d)$ is differentiable in c for each discrete choice d . Let $F(c) = \max_{d \in D} v(c, d)$ 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 $v(\cdot, \hat{d})$, 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 $\{v(\cdot, d)\}_{d \in D}$ is equi-differentiable and has uniformly bounded derivatives. Their redundant conditions conflict with Inada conditions. In contrast,

¹ Specifically, Rockafellar and Wets (1998, Proposition 8.5) establish that a function is sub-differentiable if and only if it has a differentiable lower support function. (See Clausen and Strub (2016, Appendix F) for a simpler proof.) It is then straightforward to show that a function is differentiable if and only if it is both sub- and superdifferentiable; see Kruger (2003).

our Differentiable Sandwich Lemma is applicable to problems with discrete choices and Inada conditions.²

Our second and most novel 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 at \bar{c} . Similarly, G is differentiable at \bar{c} .

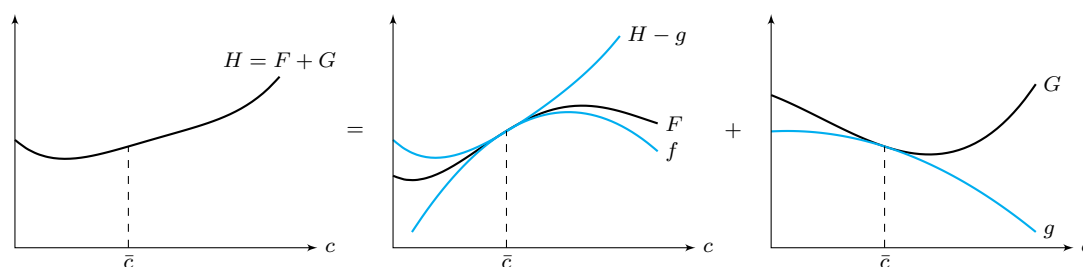


Figure 2: Reverse calculus: F is differentiable at \bar{c} .

Reverse calculus is well-suited to studying nested optimization problems in which the objective involves the solutions to other optimization problems. In these problems, 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 nested solution 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.

We combine these two tools into a **recipe for deriving first-order conditions in nested optimization problems**. Previous tools either focused on establishing the differentiability of value functions (as in the envelope theorems of Mirman and Zilcha (1975); Benveniste and Scheinkman (1979); Milgrom and Segal (2002)), or of policy functions in smooth concave problems (Araújo and Scheinkman (1977); Santos (1991)). The first half of the recipe applies the Differentiable Sandwich Lemma to establish that the objective is differentiable at optimal choices. The second half of the recipe applies Reverse Calculus to establish that the nested solution functions are differentiable at optimal choices. The main ingredient re-

² Our result accommodating Inada conditions has been applied both in computational research. Fella (2014) uses our Theorem 5 to construct a value function interpolation algorithm. Fella’s algorithm has been adapted to higher dimensions by Druedahl and Jørgensen (2017).

quired for using the recipe is the collection of differentiable upper or lower support functions, one for each nested solution function.

How are these support functions to be found? One approach is based on “lazy decision maker” constructions that involve unreactive policy functions. [Benveniste and Scheinkman’s \(1979\)](#) applied this strategy to construct a differentiable lower support function for value functions. This is a special case of a more general approach of constructing optimistic or pessimistic valuations. We construct support functions for optimal default rules based on a pessimistic value of honoring debts. Similarly, we construct a support function for a Stackelberg follower policy based on pessimistic beliefs about when the follower is capacity constrained.

Outline. [Section 2](#) formally specifies and proves the lemmas used in the recipe. [Section 3.1](#) describes our recipe for deriving first-order conditions using a Stackelberg duopoly as a running example. [Section 3.2](#) applies the recipe to Arellano’s model of unsecured credit markets. [Section 3.3](#) provides an elementary proof of the [Benveniste and Scheinkman \(1979\)](#) envelope theorem for convex dynamic programming problems. [Section 3.4](#) uses our recipe to establish first-order conditions in non-convex stochastic dynamic programming problems. [Section 4](#) concludes. [Appendix A](#) presents Reverse Calculus rules for convex combinations and function composition.

2 Techniques

This section presents the techniques used in our recipe, namely the Differentiable Sandwich Lemma and Reverse Calculus, as well as a lemma for constructing the top half of sandwiches.

2.1 Differentiable Sandwich Lemma

The Differentiable Sandwich Lemma establishes that if a function $F : C \rightarrow \mathbb{R}$ is sandwiched at some point \bar{c} between two differentiable functions, then F is differentiable at \bar{c} . We assume the domain C is a subset of a Banach space, usually \mathbb{R}^n .

Lemma 1 (Differentiable Sandwich Lemma). *Let $N \subseteq C$ be an open neighbourhood of \bar{c} . If*

- (i) $U(\bar{c}) = F(\bar{c}) = L(\bar{c})$,
- (ii) $U(c) \geq F(c) \geq L(c)$ for all $c \in N$, and
- (iii) L and U are differentiable 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 locally 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 with $\bar{c} + \Delta c \in N$,

$$\begin{aligned} & \frac{L(\bar{c} + \Delta c) - F(\bar{c}) - m \Delta c}{\|\Delta c\|} \\ & \leq \frac{F(\bar{c} + \Delta c) - F(\bar{c}) - m \Delta c}{\|\Delta c\|} \leq \frac{U(\bar{c} + \Delta c) - F(\bar{c}) - m \Delta c}{\|\Delta c\|}. \end{aligned} \quad (1)$$

Consider the limits as $\Delta c \rightarrow 0$. 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 $F'(\bar{c}) = m$. \square

2.2 Maximum Lemma

In optimization problems, the top half of the sandwich can be constructed as a horizontal line (or hyperplane) through the maximum (see [Figure 3c](#)).

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

2.3 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 A](#). We also omit the corresponding rules for subtraction and division, which involve both upper and lower support functions.

Lemma 3 (Reverse Calculus). *Suppose $F : C \rightarrow \mathbb{R}$, and $G : C \rightarrow \mathbb{R}$ 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 $H(c) = J(F(c))$ is differentiable at \bar{c} and $J : \mathbb{R} \rightarrow \mathbb{R}$ is continuously differentiable at $F(\bar{c})$ 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. \square

3 Applications

3.1 Stackelberg Duopoly

This section illustrates our recipe for deriving first-order conditions in a model of Stackelberg duopoly. The model is very simple, so the reader can focus on understanding the recipe.

In Stackelberg duopoly games, the leader's first-order condition involves the derivative of the follower's policy function. But this first-order condition is only valid if the follower's policy is differentiable (at the leader's optimal quantity). Other related problems are explored by Kydland and Prescott (1977) and Ljungqvist and Sargent (2012, Chapter 19).

We study a Stackelberg duopoly in which the follower is capacity constrained, and the leader does not know the follower's cost function. The leader chooses her output y_1 first. Then the follower chooses his output y_2 , subject to an exogenous capacity constraint $y_2 \leq Y$. The leader pays a production cost of $C(y_1, a)$. The follower is either productive or unproductive, and pays a production cost of $C(y_2, b)$ or $C(y_2, c)$. The follower knows his own cost function $z \in \{b, c\}$, but the leader only knows the probability p_z of the function $C(\cdot, z)$ occurring. The output is sold at the market price, $P(y_1 + y_2)$. If firm i has the cost function C_z , he earns a profit of $\pi(y_1, y_2, i, z) = y_i P(y_1 + y_2) - C(y_i, z)$. The follower chooses $y_2 = f(y_1, z)$ by solving

$$f(y_1, z) = \arg \max_{y_2} \pi(y_1, y_2, 2, z) \quad (2)$$

and the leader chooses y_1 to maximise

$$\phi(y_1) = \sum_{z \in \{b, c\}} p_z \pi(y_1, f(y_1, z), 1, a) = \sum_{z \in \{b, c\}} p_z y_1 P(y_1 + f(y_1, z)) - C(y_1, a). \quad (3)$$

Can we write down necessary first-order conditions for these choices? If so, the follower's and leader's first-order conditions would be

$$P(y_1 + y_2) + P'(y_1 + y_2)y_2 = C_{y_2}(y_2, z) \text{ and} \quad (4)$$

$$\sum_{z \in \{b, c\}} p_z [P(y_1 + f(y_1, z)) + P'(y_1 + f(y_1, z))(1 + f_{y_1}(y_1, z))y_1] = C_{y_1}(y_1, a). \quad (5)$$

But do the derivatives $f_{y_1}(y_1, z)$ of the follower's policies exist? We can not simply assume the follower chooses differentiable policies. In fact, the follower's policies have kinks, depicted in Figure 3a. These kinks are reflected in the leader's objective, depicted in Figure 3b.

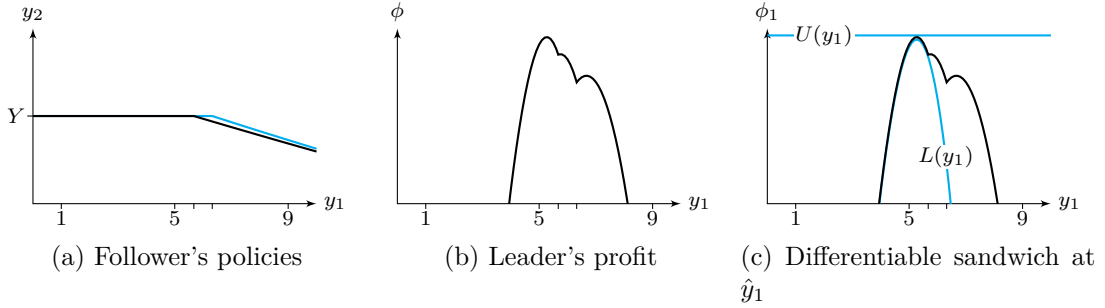


Figure 3: A Stackelberg duopoly

Theorem 1. *Assume the demand function P and cost functions $C(\cdot, z)$ are twice differentiable, and strictly concave and convex respectively. If (\hat{y}_1, \hat{y}_2) are equilibrium quantities, then (i) ϕ , $f(\cdot, b)$ and $f(\cdot, c)$ are differentiable at \hat{y}_1 , and (ii) \hat{y}_1 satisfies (5).*

Proof. The proof is a simple example of the following recipe used in all of the applications.

Ingredients: Differentiable support functions. We construct appropriate differential support functions at \hat{y}_1 for all of the functions we would like to differentiate, i.e. the leader's objective function, and the follower's policy functions.

First, we construct a differential upper support function $F(\cdot, b)$ for $f(\cdot, b)$ at \hat{y}_1 . Now, $f(\cdot, b)$ is the lower envelope of the constant function Y and the unconstrained best response function. The former is trivially differentiable, and we imposed assumptions above to ensure the latter is differentiable. The lower envelope of differentiable functions has a differentiable support function for every point.

A similar argument applies to construct a differentiable upper support function $F(\cdot, c)$ for $f(\cdot, c)$ at \hat{y}_2 .

Second, we construct a differential lower support function L for the leader's objective ϕ at \hat{y}_1 . Since $F(\cdot, z)$ involves a more aggressive best response than f_z , it involves lower profits to the leader. Therefore $L(y_1) = \sum_{z \in \{b, c\}} p_z y_1 P(y_1 + F(y_1, z)) - C(y_1, a)$ is a differentiable lower support function for ϕ .

Step 1: Upper support function. The constant function $U(y_1) = \phi(\hat{y}_1)$ is a differentiable upper support function for ϕ at \hat{y}_1 .

Step 2: Apply the Differentiable Sandwich Lemma. The support functions L and U form a sandwich around ϕ , as 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 $\phi'(\hat{y}_1) = U'(\hat{y}_1) = 0$.

Step 3: Apply Reverse Calculus. However, we have not yet established (5), which is a more useful first-order condition. In particular, we have not yet determined whether the policies $f(\cdot, b)$ and $f(\cdot, c)$ are differentiable at \hat{y}_1 . This is a Reverse Calculus problem: we have established that the left side of (3) is differentiable, and we would now like to infer that $f(\cdot, b)$ and $f(\cdot, c)$ on the right side are differentiable.

First, we apply the reverse summation rule from [Lemma 3](#). The right side of (3) is the sum of three terms, each of which has a differentiable lower support function. (The first of these functions is $y_1 \mapsto p_b y_1 P(y_1 + F(y_1, b))$.) The rule then implies that each term is differentiable at \hat{y}_1 .

We apply Reverse Calculus rules from [Lemma 3](#) as follows:

- By the reverse product rule, $P(y_1 + f(y_1, b))$ is differentiable at \hat{y}_1 .
- By the reverse chain rule, $y_1 + f(y_1, b)$ is differentiable at \hat{y}_1 .
- By the reverse summation rule, $f(\cdot, b)$ is differentiable at \hat{y}_1 .

A similar procedure establishes that $f(\cdot, c)$ is differentiable at \hat{y}_1 . □

The proof tackles two questions in turn:

- (i) Can the leader's problem be solved with a first-order condition?
- (ii) Can the leader's problem be solved with the first-order condition in (5)?

The first question is answered affirmatively if it is possible to find appropriate ingredients for the recipe. These are: a differentiable lower support function L for the objective ϕ at the solution \hat{y}_1 , and, en route, a differentiable support function for the follower's policy too. If an appropriate L can be found, then the solution satisfies a first-order condition, namely $L_1(\hat{y}_1) = 0$. The challenge is to find a pessimistic version of the leader's objective that is differentiable, or at least, differentiable at the solution \hat{y}_1 . Our approach was to consider a leader that pessimistically overestimates the follower's production, either by wrongly believing that the follower is unconstrained, or by wrongly believing that the follower never reduces production below the capacity constraint. Both types of pessimistic leader have a smooth objective, because the issue of whether or not the capacity constraint binds has been removed. The existence of a smooth pessimistic objective L establishes that even though the objective ϕ has kinks, these kinks do not create implicit boundaries where first-order conditions might fail.

The rest of the recipe is devoted to constructing more useful first-order conditions. These steps are typically less challenging than preparing the ingredients. The first two steps apply the Differentiable Sandwich Lemma to establish the first-order condition $\phi_1(\hat{y}_1) = 0$ holds at \hat{y}_1 . Then Reverse Calculus reuses the ingredient support functions to unpack this first-order condition into (5).

3.2 Unsecured Credit

Our second application is about unsecured debt contracts where borrowers may decide either to 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. 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 sizes.

All four threats to first-order conditions from the introduction are present. First, the borrower's future default policy appears in his objective, because it determines default risk and hence interest rates. 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, we impose (problematic) Inada conditions so that the borrower makes interior consumption choices.

We apply our recipe to establish that both nested solution functions – the value function and the interest rate – are differentiable at optimal debt choices (except at the risk-free credit limit). Hence, we derive a first-order condition involving a marginal interest rate and a marginal continuation value. We also apply Reverse Calculus to express the marginal interest rate in terms of the cut-off policy; this allows us to express the first-order condition in terms of quantities rather than prices. Finally, we apply our recipe to characterise the borrower’s credit limit and establish that this limit is never exhausted.

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. One paper by [Aguiar, Amador, Hopenhayn and Werning \(2019\)](#) derives first-order conditions for a simple example, where it is possible to avoid differentiating any policy functions. The following three papers apply some Euler equations, with the first explicitly acknowledging that they lack justification for differentiating the interest rates with respect to loan size. We provide a justification. [Aguiar and Gopinath \(2006\)](#) dropped a detailed discussion of their heuristic Euler equation from their NBER working paper version. Similarly, [Arellano and Ramanarayanan \(2012\)](#) use heuristic Euler equations to compare maturity structures of loans. Finally, [Hatchondo and Martinez \(2009\)](#) discuss an Euler equation, implicitly assuming differentiability of interest rates.

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. $\lim_{c \rightarrow 0^+} u_1(c) = \infty$. 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 promises to pay a lender b' in the following period, although both understand that the borrower only has an incentive to honour the promise if tomorrow’s x' lies in some set H' . Thus, a debt contract consists of (b', H') . 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 – regardless of whether $x' \in H'$ – 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.³ The borrower's autarky value after defaulting is

$$A(x) = u(x - s) + \beta \int A(x')f(x')dx'. \quad (6)$$

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 $V(b', x')$, should not be less than the autarky value $A(x)$. The borrower's value of honouring debts is therefore⁴

$$\begin{aligned} V(b, x) = & \sup_{c, b', H'} u(c) + \beta \int \max \{A(x'), V(b', x')\} f(x')dx', \\ \text{s.t. } & c + b = x + \left[\beta \int_{H'} f(x') dx' \right] b', \\ & V(b', x') \geq A(x') \text{ for all } x' \in H', \\ & b' \leq b^{\text{ponzi}}. \end{aligned} \quad (7)$$

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 $(b', H') = (b', [y(b'), x^{\text{max}}])$, 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

$$V(b, x) = \sup_{b' \leq b^{\text{ponzi}}} \phi(b'; b, x), \quad (8)$$

³ 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.

⁴ We mention some technicalities: (i) the borrower should be constrained to choosing a measurable honour set, and (ii) since we focus on first-order conditions, we take it for granted that the value function in the sequence problem is the unique solution to the Bellman equation.

where

$$\phi(b'; b, x) = u(x - b + q(b')b') + \beta W(b'), \quad (9a)$$

$$W(b') = \int \max \{A(x'), V(b', x')\} f(x') dx', \quad (9b)$$

$$q(b') = \beta [1 - F(y(b'))], \quad (9c)$$

$$y(b') = \min (\{x' \in [x^{\min}, x^{\max}] : V(b', x') \geq A(x')\} \cup \{x^{\max}\}). \quad (9d)$$

We denote optimal policy functions by $\hat{b}'(b, x)$.⁵

The objective (9a) has two nested solution functions, q and W , which we will show are not globally differentiable. The value function W has downward kinks at states of indifference between honouring and defaulting, as depicted in Figure 1c. Similarly, we have no a priori knowledge of the differentiability of the debt price q .

We will follow the four steps of the recipe to establish that at optimal choices, first-order conditions hold and that $\phi(\cdot, b, x)$, q and W are differentiable. However, we find that there is one exception: the debt price exhibits an upward kink at the risk-free credit limit. This means that first-order conditions are inapplicable when the borrower chooses to exhaust his risk-free credit limit.

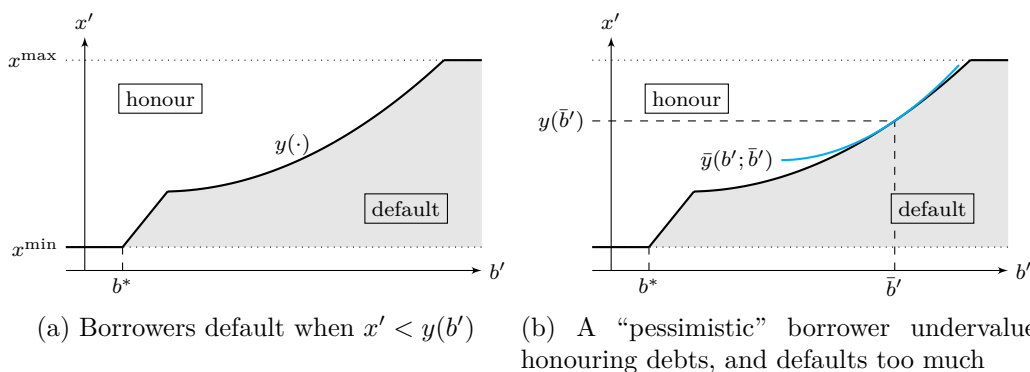


Figure 4: The default cut-off rule

Ingredients: Differentiable Lower Support Functions. The most important ingredient of the recipe is providing appropriate differentiable support functions for the nested solution 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

⁵ The borrower might be indifferent between several optimal policies.

Figure 4a. 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 pessimistic borrower that undervalues honouring debts, and hence uses a higher cut-off than $y(\cdot)$. Specifically we consider a pessimistic 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 pessimistic borrower's value function is

$$L(b', x'; \bar{b}') = u(x' - b' + q(\bar{b}'')\bar{b}'') + \beta W(\bar{b}''). \quad (10)$$

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

$$L(b', \bar{y}(b'; \bar{b}'); \bar{b}') = A(\bar{y}(b'; \bar{b}')) \text{ for all } b' \quad (11)$$

provides an upper support function for the cut-off $y(\cdot)$ at \bar{b}' , depicted in **Figure 4b**. Since the pessimistic 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 $\bar{b}' > b^*$.⁶

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 4. *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 $V(b', x')$. However, this time, we use a lazy borrower's value function that differs from the pessimistic value function used to construct (10). 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

$$M(b', x'; \bar{b}') = u(x' - b' + q(\bar{b}''(x'))\bar{b}''(x')) + \beta W(\bar{b}''(x')). \quad (12)$$

⁶ Applying the implicit function theorem to (11) gives

$$\bar{y}_1(\bar{b}'; \bar{b}') = \frac{u_1(\bar{c}'(\bar{b}', y(\bar{b}')))}{u_1(\bar{c}'(\bar{b}', y(\bar{b}')) - u_1(y(\bar{b}') - s)} > 1.$$

This means that,

$$\underline{W}(b'; \bar{b}') = A(x') + \int_{\bar{y}(b'; \bar{b}')}^{x^{\max}} [M(b', x'; \bar{b}') - A(x')] f(x') dx' \quad (13)$$

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.⁷ 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 $b' = \bar{b}'$ with⁸

$$\underline{W}_1(b'; \bar{b}') = \int_{\bar{y}(b'; \bar{b}')}^{x^{\max}} M_1(\bar{b}', x'; \bar{b}') f(x') dx'. \quad (14)$$

This means we have established:

Lemma 5. *At every \bar{b}' , there exists a differentiable lower support function $\underline{W}(\cdot; \bar{b}')$ for W .*

Combining the differentiable support functions for q and W from [Lemma 5](#) and [Lemma 4](#), we conclude that for all (b, x) and all $\bar{b}' \neq b^*$, the function

$$b' \mapsto u(x - b + q(b'; \bar{b}')b') + \beta \underline{W}(b'; \bar{b}') \quad (15)$$

is a differentiable lower support function for $\phi(\cdot; b, x)$ at \bar{b}' .

We now have all the ingredients ready to apply the recipe.

Theorem 2. *Suppose $\hat{c}(\cdot, \cdot)$ and $\hat{b}'(\cdot, \cdot)$ are optimal policies for (8).⁹ Fix any state (b, x) and let $\hat{b}' = \hat{b}'(b, x)$. Then either*

- (i) $\hat{b}' = b^*$, or
- (ii) $\hat{b}' < b^*$ and \hat{b}' satisfies the first-order condition

$$u_1(\hat{c}(b, x)) = \beta W_1(\hat{b}') = \int u_1(\hat{c}(b', x')) f(x') dx', \quad (16)$$

and the value function W is differentiable at \hat{b}' , or

⁷See the Measurable Maximum Theorem in [Aliprantis and Border \(2006, Theorem 18.19\)](#).

⁸See for example [Weizsäcker \(2008, Theorem 4.6\)](#).

⁹Note that c does not appear in (8). It is shorthand for $x - b + q(b')b'$.

(iii) $\hat{b}' > b^*$ and \hat{b}' satisfies the first-order condition

$$u_1(\hat{c}(b, x))(q(\hat{b}') + q_1(\hat{b}')\hat{b}') = \beta W_1(\hat{b}'), \quad (17)$$

and the nested functions W , q and y are differentiable at \hat{b}' with

$$W_1(\hat{b}') = \int_{y(\hat{b}')}^{x^{max}} u_1(\hat{c}(\hat{b}', x')) f(x') dx', \quad (18)$$

$$q_1(\hat{b}') = -\beta F_1(y(\hat{b}')) y_1(\hat{b}'), \quad (19)$$

$$y_1(\hat{b}') = \frac{u_1(\hat{c}(\hat{b}', y(\hat{b}')))}{u_1(\hat{c}(\hat{b}', y(\hat{b}'))) - u_1(y(\hat{b}') - s)}. \quad (20)$$

Proof. The first case does not involve any first-order conditions. We apply the recipe for the other two cases.

Step 1. At the optimal choice $\hat{b}' = \hat{b}'(b, x)$, the borrower's objective $\phi(\cdot; b, x)$ has a trivial upper support function $U(b'; b, x) = \phi(\hat{b}', b, x)$.

Step 2. At \hat{b}' , the borrower's objective $\phi(\cdot; b, x)$ is sandwiched between (15) and $U(\cdot; b, x)$. So, the Differentiable Sandwich Lemma (Lemma 1) implies the borrower's objective is differentiable at the optimal debt choice \hat{b}' .

Step 3. Repeated application of Reverse Calculus (Lemma 3) implies that W (in the second case) or W , q and y (in the third case) are differentiable at \hat{b}' .¹⁰ \square

The first-order condition (17) can be interpreted as follows. The borrower equates the marginal benefit of owing debt with the marginal cost. 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). The marginal benefit consists of the marginal utility of consumption times the marginal revenue from promising an extra payment to the lender. Reverse calculus allows us to quantify the marginal revenue of promises. Specifically, when promising an

¹⁰ We repeatedly apply Lemma 3 as follows. First, we apply rule (i) (summation) to (9a) to establish that both terms, $b' \mapsto u(x - b + q(b')b')$ and $b' \mapsto \beta W(b')$ are differentiable at \hat{b}' . Hence W is differentiable at \hat{b}' . If $\hat{b}' < b^*$, then we know that $q(b') = \beta$, so we stop here, giving the second case. Otherwise we are in the third case, and we continue. We apply rule (iv) (function composition) to $b' \mapsto u(x - b + q(b')b')$, which establishes that $b' \mapsto q(b')b'$ is differentiable at \hat{b}' . Next, we apply rule (ii) (multiplication) to establish that q is differentiable at \hat{b}' . This also means that both sides of (9c) are differentiable. So we apply rule (iv) (function composition) to the right side, $b' \mapsto \beta[1 - F(y(b'))]$ and conclude that y is differentiable at \hat{b}' .

extra payment, the honour probability decreases according to q_1 , which reflects an increase in the default cut-off of y_1 .

Also note that the theorem gives formulae for all derivatives. In particular all prices and marginal prices can be written in terms of quantities. This means it is possible to write the first-order condition in terms of quantities only, which can be helpful in computational work.¹¹

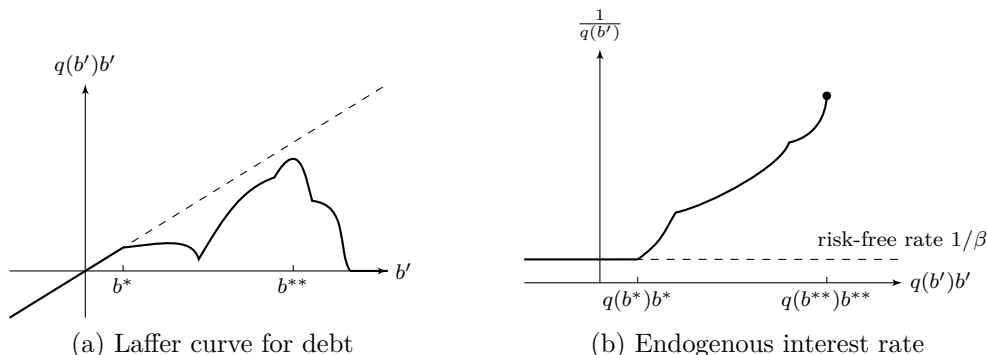


Figure 5: Characterisation of endogenous borrowing

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 5a. The borrower's credit limit is the maximum of this curve, $q(b**)b**$, where

$$b** = \arg \max_{b'} q(b')b'. \quad (21)$$

We apply the recipe for this new optimisation problem. If $b** > b^*$, then we have already constructed a differentiable lower support function for q at $b**$, so the Differentiable Sandwich Lemma (Lemma 1) together with the Reverse Calculus Lemma (Lemma 3) imply that q is differentiable at $b**$ with

$$q(b**) + q_1(b**)b** = 0. \quad (22)$$

Substituting this into the Euler equation (17), we see that the marginal benefit of taking on debt at $b**$ is zero, while the marginal cost is positive. Therefore, we conclude

¹¹ A FOC without prices is $u_1(\hat{c}(b, x))[1 - F(y(\hat{b}')) - F_1(y(\hat{b}'))y_1(\hat{b}')\hat{b}'] = W_1(\hat{b}')$.

Theorem 3. *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 5b](#). 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 5a](#). 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. Our main contributions here are: (i) establishing that first-order conditions – involving derivatives of both policy functions and value functions – hold at optimal choices, (ii) providing formulas for these derivatives, so (marginal) prices can be written in terms of quantities, and (iii) applying the same logic to a Laffer curve for debt, concluding that borrowers do not exhaust their credit limits.

Others have applied these techniques to study various issues in both consumer and sovereign credit markets; see [Jeske, Krueger and Mitman \(2013\)](#) and [Müller, Storesletten and Zilibotti \(2016\)](#).

To keep the analysis simple, we assumed that shocks are IID. This condition 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.

There are two potential ways to accommodate persistent shocks. First, [Chatterjee, Corbae, Nakajima and Ríos-Rull \(2007, Theorem 3\)](#) established that two-sided cut-off rules are optimal in an environment with persistent shocks. 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.

Second, we could assume that all persistence is summarised by a finite Markov chain. The shock x would be drawn from a distribution determined by a Markov state, such as “boom” or “recession”. There would be one cut-off rule for each Markov state, and hence one risk-free credit limit for each state. Hence there would be multiple debt-levels where first-order conditions are inapplicable.

3.3 Convex Dynamic Programming

Benveniste and Scheinkman (1979) study *value* functions in smooth convex 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:

$$V(c) = \sup_{c' \in \{\tilde{c}' : (c, \tilde{c}') \in \Gamma\}} u(c, c') + \beta V(c'), \quad (23)$$

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.

Theorem 4 (Benveniste-Scheinkman Theorem). *If \tilde{c}' is an optimal choice at state $c \in \text{int}(\{\tilde{c} : (\tilde{c}, \tilde{c}') \in \Gamma\})$, then V is differentiable at c with $V_c(c) = u_c(c, \tilde{c}')$.*

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 $L(c) = u(c, \tilde{c}') + \beta V(\tilde{c}')$. Lemma 1 delivers the conclusions. \square

Graduate economics textbooks such as Stokey and Lucas (1989) do not provide a self-contained proof of this theorem. 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.

3.4 Non-Convex Dynamic Programming

In this section, we study a more general class of stochastic dynamic programming problems. We drop all convexity assumptions, and we accommodate discrete choice sets and uncertainty. As discussed in the introduction, discrete choices can lead to kinks, and the expectations operator can hide problematic kinks. Nevertheless, we establish that recursive first-order conditions hold at optimal choices.

Problem 2. Each period, a household chooses $(c', d') \in \Omega$, where $c' \in \mathbb{R}^n$ are *continuous* choices and d' are *discrete* choices. The state variable consists of the previous period's choices (c, d) as well as an exogenous Markov state $\theta \in \Theta$ with transition matrix π . We assume that the per-period utility $u(c, c'; d, d'; \theta)$ is differentiable with respect to c and c' . The set of combinations of states and feasible choices is denoted Γ . The Bellman equation is

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

s.t. $(c, c'; d, d'; \theta) \in \Gamma$,

where the domain of V is $\Omega \times \Theta$.

Suppose that $\omega(c, d, \theta)$ is an optimal choice vector at the state (c, d, θ) .

Definition 1. Fix any state (c, d, θ) , and consider the optimal choices $(\bar{c}', \bar{d}') = \omega(c, d, \theta)$ and $(\bar{c}''(\theta'), \bar{d}''(\theta')) = \omega(\bar{c}', \bar{d}', \theta')$ for the following two periods. Then the set of feasible one-shot deviations from \bar{c}' is

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

Theorem 5. Let $(\hat{c}', \hat{d}') = \omega(c, d, \theta)$ 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 $(\hat{c}''(\theta'), \hat{d}''(\theta')) = \omega(\hat{c}', \hat{d}', \theta')$.

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

$$\phi(c') = u(c, c'; d, \hat{d}'; \theta) + \beta \sum_{\theta' \in \Theta} \pi_{\theta\theta'} V(c', \hat{d}', \theta'), \quad (24)$$

where the domain of ϕ is $\Lambda(c, d, \theta)$.

Ingredients. We prepare the ingredient support functions. V has a differentiable lower support function at $(\hat{c}', \hat{d}', \theta')$,

$$v(c', \hat{d}', \theta') = u(c', \hat{c}''(\theta'); d', \hat{d}''(\theta'), \theta') + \beta \sum_{\theta'' \in \Theta} \pi_{\theta'\theta''} V(\hat{c}''(\theta'), \hat{d}''(\theta'), \theta''). \quad (25)$$

This leads to a differentiable lower support function for ϕ at \hat{c}' , namely

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

Since we assumed \hat{c}' is an interior choice, v and L are well-defined in an open neighbourhood of \hat{c}' .

Step 1. ϕ has a differentiable upper support function at \hat{c}' , namely $U(c') = \phi(\hat{c}')$.

Step 2. ϕ is sandwiched between L and U at \hat{c}' . By the Differentiable Sandwich Lemma, ϕ is differentiable at \hat{c}' with $\phi_{c'}(\hat{c}') = U_{c'}(\hat{c}') = 0$.

Step 3. The addition rule of Reverse Calculus implies that V is differentiable with respect to c' at each $(\hat{c}', \hat{d}', \theta')$. \square

Milgrom and Segal (2002, Corollary 2) proved a special case of this theorem in a static context (when $\beta = 0$), and with two redundant assumptions, namely that the utility function is equidifferentiable and the marginal utilities are uniformly bounded. Milgrom and Segal make good use of these assumptions in establishing global differentiability properties of value functions in their Theorem 3. But these assumptions are redundant for studying first-order conditions.

Both conditions are problematic. The equidifferentiability condition is difficult to check in infinite horizon dynamic problems. More importantly, the uniformly bounded derivative condition conflicts with Inada conditions. Inada conditions are often imposed to ensure first-order conditions hold by directing optimal choices away from boundaries.

Milgrom and Segal (2002) impose these conditions in order to ensure that the value function's directional derivatives exist globally. Similarly, other papers make assumptions including Lipschitz continuity (Clarke (1975)) and supermodularity (Amir, Mirman and Perkins (1991)) to ensure the existence of directional derivatives. Our recipe does not make use of directional derivatives, so we can dispense with these assumptions.

4 Conclusion

We can not assume that the solutions to nested optimization problems are smooth. Nevertheless, we found that in some economic problems, it is optimal for decision makers to choose from smooth parts of their menus. Nested optimization problems come in many forms, so we do not provide a one-size-fits-all theorem. Instead, we

devised a recipe based on the Differentiable Sandwich Lemma. Applying the recipe requires finding appropriate differentiable support functions.

There are potentially many ways to mix and match 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) pessimistic cut-off rules. Of these constructions, only the Reverse Calculus construction is truly unprecedented. The power of our approach derives from the ability to combine these constructions, and the three-step recipe provides an intuitive way to organise them. 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.

A 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 6. *Suppose $E : C \rightarrow [0, 1]$, $F : C \rightarrow \mathbb{R}$, and $G : C \rightarrow \mathbb{R}$ 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 differentially sandwiched between h and H at \bar{c} . By the Differentiable Sandwich Lemma, \tilde{h} is 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} . \square

Finally, we consider function composition of two endogenous functions.

Lemma 7. *If $H(c) = J(K(c))$ is differentiable at \bar{c} , where*

- $J : \mathbb{R} \rightarrow \mathbb{R}$ has an inverse J^{-1} and a differentiable lower support function $j(\cdot)$ at $K(\bar{c})$,
- $K : \mathbb{R} \rightarrow \mathbb{R}$ 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 $j'(K(\bar{c})) > 0$.¹² 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. \square

¹² 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 $\tilde{j}'(c) = \frac{1}{[j(c)+x_0]^2}j'(c)$.

References

- AGUIAR, M., AMADOR, M., HOPENHAYN, H. and WERNING, I. (2019). Take the short route: Equilibrium default and debt maturity. *Econometrica*, **87**, 423–462.
- 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.
- and RAMANARAYANAN, A. (2012). Default and the maturity structure in sovereign bonds. *Journal of Political Economy*, **120** (2), 187–232.
- 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.
- BULOW, J. and ROGOFF, K. (1989). Sovereign debt: Is to forgive to forget? *American Economic Review*, **79** (1), 43–50.
- 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.
- CLAUSEN, A. and STRUB, C. (2016). Introduction to economic theory, lecture notes.
- DRUEDAHL, J. and JØRGENSEN, T. H. (2017). A general endogenous grid method for multi-dimensional models with non-convexities and constraints. *Journal of Economic Dynamics and Control*, **74**, 87 – 107.
- EATON, J. and GERSOVITZ, M. (1981). Debt with potential repudiation: Theoretical and empirical analysis. *Review of Economic Studies*, **48** (2), 289–309.
- FELLA, G. (2014). A generalized endogenous grid method for non-concave problems. *Review of Economic Dynamics*, **17** (2), 329–344.
- HATCHONDO, J. C. and MARTINEZ, L. (2009). Long-duration bonds and sovereign defaults. *Journal of International Economics*, **79** (1), 117–125.
- JESKE, K., KRUEGER, D. and MITMAN, K. (2013). Housing and the macroeconomy: The role of bailout guarantees for government sponsored enterprises. *Journal of Monetary Economics*, **60**, 409–442.
- 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.

- 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.
- MÜLLER, A., STORESLETTEN, K. and ZILIBOTTI, F. (2016). Sovereign debt and structural reforms.
- ROCKAFELLAR, R. T. (1970). *Convex Analysis*. Princeton University Press.
- and WETS, R. J. (1998). *Variational Analysis*. Springer.
- SANTOS, M. S. (1991). Smoothness of the policy function in discrete time economic models. *Econometrica*, **59** (5), 1365–1382.
- STOKEY, N. L. and LUCAS, R. E. (1989). *Recursive Methods in Economic Dynamics*. Harvard University Press.
- WEIZSÄCKER, H. v. (2008). Basic measure theory, lecture notes.