

The marginal utility of money:
A modern Marshallian approach to consumer choice*

Daniel Friedman

József Sákovics

University of California at Santa Cruz

The University of Edinburgh

September 8, 2011

Abstract

We reformulate neoclassical consumer choice by focusing on λ , the marginal utility of money. As the opportunity cost of current expenditure, λ is approximated by the slope of the indirect utility function of the continuation. We argue that λ can largely supplant the role of an arbitrary budget constraint in partial equilibrium analysis. The result is a better grounded, more flexible and more intuitive approach to consumer choice.

JEL classification: D01, D03, D11.

Keywords: budget constraint, separability, value for money.

*We thank Luciano Andreozzi, William Brock, Joan-Maria Esteban, Steffen Huck, Axel Leijonhufvud, Youn Kim, Michael Mandler, Carmen Matutes, David de Meza, Ryan Oprea, Martin Shubik, Nirvikar Singh, Hal Varian, Donald Wittman and other June 2011 UCSC brown bag seminar participants for helpful suggestions, and thank Olga Rud for help in preparing the figures.

1 Introduction

When faced with the simple task of deciding how much to buy of a particular good at a given price, *Homo Economicus* solves a horrendously complex problem. She maximizes her utility by choosing a lifetime consumption plan that takes into consideration all the contingencies she might face in the immediate and distant future (as well as frictions, uncertainty, and other complications), and then executes the first component of this plan by purchasing the optimal amount of the good in question.

Not surprisingly, economists have found ways to make consumer choice theory more tractable, applying a “partial equilibrium” approach. The textbook approach assumes, explicitly or implicitly, that the good in question has at most a few substitutes and complements, and chops off this small portion of the lifetime problem. To this portion it applies a budget constraint, and analyzes the budget-constrained consumer’s reaction to surprises in prices and availability of the target good and related goods, as well as to unplanned changes in the budget itself (e.g., Hal Varian, 1992, Ch. 7-8).

Unfortunately the textbook analysis has serious shortcomings. By fully separating the subproblem from the larger problem, the budget constraint rules out substitution of purchasing power across the boundary, irrespective of the realization of prices, etc. As we will see, a fixed budget constraint distorts the solution because the size of the optimal budget is highly sensitive to variations in the subproblem. One could mitigate the distortion by broadening the subproblem – perhaps to a set where a *bona fide* liquidity constraint binds – but that would forfeit the simplicity of partial equilibrium and would tie together decisions over goods whose consumption utilities are independent of each other. Thus standard partial equilibrium analysis does violence to the underlying general equilibrium problem, but it is widely accepted that this is the price that must be paid for tractability.

This essay will challenge that view. We argue for a robust, flexible and natural approach to disaggregation, simpler than the textbook approach but no less rigorous. The basic idea is to use the *marginal utility of money*, rather than the budget constraint, to link the subproblem to the rest-of-life problem. Using a cardinal utility function, the consumer can react to changes in the environment by substituting optimally within the subproblem and also by optimally shifting purchasing power across the subproblem boundary. The marginal

utility of money provides a robust criterion for the trade-off between subproblem and rest-of-life, and it resonates with the findings of consumer research.

Section 2 presents a standard general equilibrium formulation of the consumer's lifetime problem, and offers a simple definition of separable subproblems. After reviewing the textbook approach to the subproblem, it shows how our new Marshallian approach offers a somewhat different solution. The next section compares textbook comparative statics to those of the new Marshallian approach: the substitution effect is the same, but income effects and overall effects differ.

Section 4 then extends the new Marshallian approach to indivisible goods, and to larger subproblems. It shows how the consumer can use personal experience to update the marginal utility of money without trying to resolve the entire lifetime problem. The section also shows how to model liquidity-constrained consumers, including those who live paycheck to paycheck.

Section 5 discusses further applications of the new Marshallian approach. The marginal utility of money provides a simple heuristic that seems consistent with descriptions of consumer behavior in the marketing literature, and is a useful counterpoint to the behavioral economics notion of mental accounting. Somewhat more speculatively, the section proposes connections to the management of multidivisional firms and to money illusion.

A concluding discussion reiterates that, compared to the textbook approach, the new Marshallian approach offers (a) more robust prescriptions for how consumers should react to surprises, (b) a better way to connect partial equilibrium to general equilibrium analysis, and also (c) more plausible descriptions of actual human behavior.

An historical perspective may be helpful before we begin. The idea of a cardinal utility function defined over purchasing power goes back at least to Jeremy Bentham (1802), as does the argument that marginal utility diminishes. Alfred Marshall (1890, 1920), synthesizing the work of earlier Marginalists, obtained the crucial first order condition (that marginal utility for each good equals its price times the marginal utility of money) in the special case that total utility is additively separable in each good. "Edgeworth destroyed this pleasant simplicity and specificity when he wrote the total utility function as $f(x_1, x_2, x_3, \dots)$," says George Stigler (1950, p. 322). It fell to John Hicks and Roy Allen (1934) to show how to

impose a budget set to derive demand functions and cross-price elasticities when goods might have complements and substitutes. Their analysis, developed further by Paul Samuelson and a host of other economists, eventually became textbook orthodoxy.

We present a model in the spirit of Marshall that can deal with Edgeworth's complications. The model uses cardinal utility to obtain price elasticities that closely approximate the general equilibrium elasticities.

2 The consumer choice problem

We begin by showing how to separate a tractable subproblem from the horrendously complex lifetime choice problem, and then distinguish our new Marshallian solution of the subproblem from the textbook solution.

2.1 A subproblem of the lifetime problem

Let $X \in \mathfrak{R}^N$ represent an agent's lifetime plan of work and consumption. Using the notation $x_i^- = \min\{0, x_i\} \leq 0$ and $x_i^+ = x_i - x_i^- = \max\{0, x_i\} \geq 0$, work is represented by the negative components $X^- = (x_1^-, \dots, x_N^-)$ – analogous to inputs of a production function – and consumption by the positive components $X^+ = (x_1^+, \dots, x_N^+)$ – analogous to outputs, with investment in human capital included. Of course, the number N of goods is astronomically large, especially if we follow Arrow and Debreu in indexing goods separately by location, date and (for unresolved contingencies) the realized state of Nature.

The agent takes as given a price vector $P \in \mathfrak{R}_+^N$, and has preferences represented by a utility function U defined over a set that contains all feasible plans. “Feasible” means that X satisfies any relevant technological constraints (e.g., that a day's activities can be done in 24 hours) and also that it satisfies $PX = \sum_{i=1}^N P_i X_i = 0$. That is (after including special sorts of work such as selling endowed assets and special sorts of consumption such as taxes and gifts), lifetime income $-PX^- > 0$ equals lifetime expenditure PX^+ . It will be convenient to refer to exogenous transfers of purchasing power by changing the endowment of an additional good $i = 0$, whose price P_0 is normalized to 1.

Assume for now (later we will relax this) that all purchasing power is liquid: $L = x_0 - PX^- > 0$ is available without further constraint. Letting $\xi(L, P)$ denote the set of feasible plans, the indirect lifetime utility function is

$$V_o(L, P) = \max U(X) \text{ s.t. } X \in \xi(L, P). \quad (1)$$

A *consumer subproblem* is to choose an n -subvector of the consumption vector X^+ of a lifetime plan X . By rearranging the indexing, we can write the subvector as $x \in \xi^{[n]}(L, P) \subset \mathfrak{R}_+^n$ and the rest of the life plan as $X^\perp = (x_0, 0, \dots, 0, x_{n+1}, x_{n+2}, \dots, x_N) \in \xi^\perp(L, P)$, where $\xi^{[n]}(L, P)$ denotes the n -subvectors of ξ , and $\xi^\perp(L, P)$ denotes the complementary subvectors. The corresponding price subvector of P is denoted $p \in \mathfrak{R}_+^n$. Typically n is small, perhaps just 1 or 2.

The subproblem is *separable* if, for some utility function u defined on \mathfrak{R}_+^n we can, with negligible error, write

$$U(X) = u(x) + U(X^\perp) \quad (2)$$

for any feasible plan X . A sufficient condition for separability is that the cross second partial derivative $U_{ij} = 0$ everywhere for all $i = 1, \dots, n$ and $j = n + 1, \dots, N$.

Substituting equation (2) into (1), we have the following recursion:

$$V_o(L) = \max_{x \in \xi^{[n]}} \{u(x) + V(L - px)\}, \quad (3)$$

where $V(L)$ denotes the rest-of-life indirect utility function, defined as in (1) but with X restricted to $\xi^\perp(L, P)$. The dependence on the rest-of-life prices P^\perp is suppressed in order to emphasize the impact of the subproblem price vector p . The equation says that if today's subproblem is separable, then the only effect today's consumption has on subsequent utility is via today's expenditure $px = \sum_{i=1}^n p_i x_i$, which reduces the consumer's subsequent purchasing power. Of course, the subproblem need not be separated chronologically, but if it is then (3) can be regarded as a Bellman equation with discounting built into V .

Thus the consumer's subproblem is to choose her optimal basket of n goods that have possibly interdependent consumption values, but that are separable from the consumption values of all other goods. Note that the subproblem might be intertemporal in nature.

2.2 Solving the subproblem

The textbook solution method is to impose a budget $B > 0$ on the subproblem.¹ Assuming that there are no other constraints, the subproblem becomes

$$\max_{x \geq 0} u(x) \text{ s.t. } px \leq B. \quad (4)$$

To focus on relevant issues, we assume henceforth that u is twice continuously differentiable and strictly monotone increasing, and that (4) has a regular interior solution. The Lagrangian is

$$\max_{(x, \mu) \geq 0} [u(x) + \mu(B - px)], \quad (5)$$

with first-order conditions

$$\begin{aligned} u_i(x^*) &= \mu p_i, \quad i = 1, \dots, n \\ B &= px^*. \end{aligned} \quad (6)$$

By regularity, the strict second order condition also holds: the Hessian matrix $D^2u(x^*) = (u_{ij}(x^*))_{i,j=1,\dots,n}$ is negative definite on the tangent space normal to p . For $n = 2$ that condition implies that the determinant $u_{11}u_{22} - u_{12}u_{21} > 0$. The unique solution to (5) is the textbook demand function, denoted $x^*(p, B)$.

We recommend a different way to solve the subproblem. The idea is to endogenize expenditure and, instead of an arbitrary budget B , to focus on the opportunity cost of expenditure in the lifetime problem. That opportunity cost, which we refer to as the marginal utility of money, is obtained by linearizing the indirect utility function V around L . Linearization is reasonable because the range of sensible subproblem expenditures typically is dwarfed by life-time expenditure. (Section 4.2 below will discuss large subproblems for which a linear approximation is inadequate.)

For the approach to make sense, V must be *cardinal*. That is, it must be defined up to an increasing linear transformation, as in von Neumann and Morgenstern. By construction, V is increasing in L and we shall further assume that it is smooth and concave, i.e., $V'(L) > 0$

¹One seldom asks where B comes from. To find the optimal B , one has to solve the lifetime problem (3), which means that we have no simplification at all. In practice, B is apparently chosen via some unspecified rule of thumb.

and $V''(L) \leq 0$ for L in the relevant range. A sufficient condition is that, analogous to production functions, U is smooth, monotone, and exhibits decreasing returns to scale in relevant regions.

Thus a preliminary step is to take the first-order Taylor approximation of the indirect utility function for the lifetime problem. Defining $\lambda = V'(L)$ as the *marginal utility of money*, we have

$$V(L - px) \approx V(L) - \lambda px \text{ for } px \ll L. \quad (7)$$

Substituting the linearization (7) into the lifetime optimization problem (3) and dropping the constant term $V(L)$, we obtain the subproblem objective function, which is subject to no further constraint

$$\max_{x \geq 0} [u(x) - \lambda xp]. \quad (8)$$

In the subproblem (8), the parameter λ is exogenous, and the equation invites alternative characterizations of its role. Thus, the marginal utility of money may also be thought of as the opportunity cost of today's expenditure, or the "shadow utility" of purchasing power, or the "conversion rate" between utility and money. Note that λ depends only on V and L , and is independent of u and p in any small subproblem. Indeed, the same value of λ applies to any subproblem, as long as the linear approximation (7) is valid. By contrast, the optimal choice of the exogenous parameter B in the standard approach is intimately tied to the characteristics of the subproblem and hence is much less robust. We will elaborate this point in Section 3.2 below.

Again assuming a regular interior solution, the first-order conditions for (8) are

$$u_i(x^*) = \lambda p_i, \quad i = 1, \dots, n. \quad (9)$$

We denote the "Marshall-Edgeworth" demand function resulting from the solution of (9) by $M(p, \lambda)$. Of course, from (9) one can still recover the usual textbook tangency condition that for any pair of the n goods in the optimal bundle, the marginal rate of substitution is equal to the price ratio. Here it is the constant marginal utility of money that gives us that result, rather than a fixed budget.

3 Comparing solutions

Observe that (9) is the same as (6) except that λ replaces μ and the budget constraint is dropped. Dropping one equation makes sense because λ is exogenous in subproblem (8). But what is the relation between the marginal utility of money, λ , and the shadow price of the budget constraint, μ ? Recall that μ is the marginal consumption utility of a dollar spent over the budget. If B is not chosen optimally for the lifetime problem (3), then μ differs from the true marginal cost of overspending, $V'(L - B) \approx V'(L) = \lambda$. Thus one can regard μ as an estimate of λ , where the quality of the estimate depends on the quality of the choice of the budget.

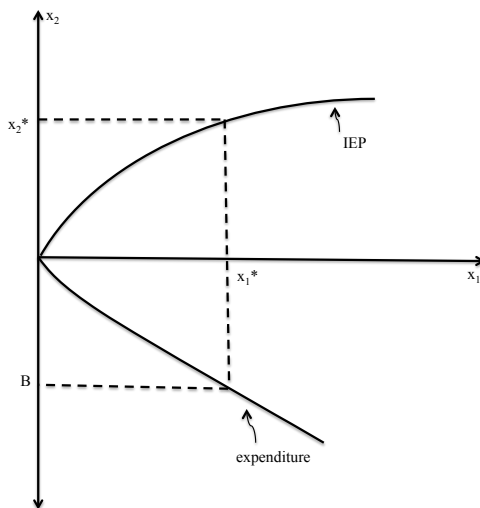


Figure 1: Optimal consumption with fixed budget B . The curve in the lower quadrant shows expenditure $e(x)$ along the IEP as a function of consumption x_1 of the first good. Given $B > 0$ on the (downward pointing) vertical axis, one finds the corresponding point on the expenditure curve to locate optimal consumption x_1^* of good 1 on the horizontal axis, and then uses the IEP in the upper quadrant to locate x_2^* on the (upward pointing) vertical axis.

The two approaches are connected via the Income Expansion Path (IEP), the locus of points that satisfy the tangency conditions (6) for a given price vector as B varies from zero to infinity. Since the tangency conditions are the same in (9), viz., that the price ratio equals the marginal rate of substitution, exactly the same locus of utility maximizing bundles is traced out by (9) as λ varies from infinity to zero.

Consequently, for any price vector p , the textbook demand function $x^*(p, B)$ and the

Marshall-Edgeworth demand function $M(p, \lambda)$ both pick out points on the *same* IEP. The earlier discussion shows that the two different decision rules will typically select different points on this IEP. Figure 1 illustrates the choice of $x^*(p, B)$, and Figure 2 also illustrates the choice of $M(p, \lambda)$. The intuition for $M(p, \lambda)$ is that the consumer moves up the IEP as long as the utility gain exceeds the opportunity cost of the expenditure, and stops when the gain diminishes to the point that it is equal to the opportunity cost.

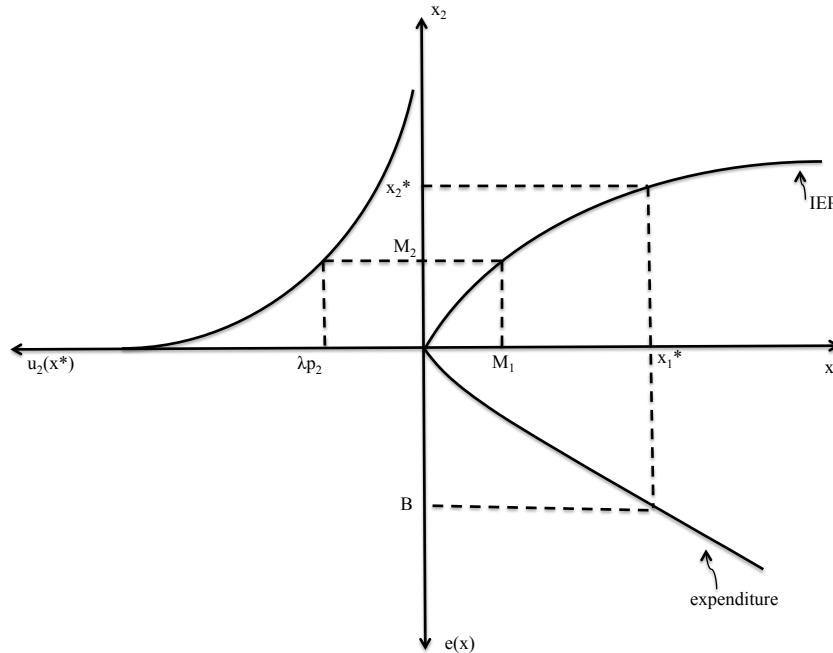


Figure 2: Optimal consumption with given λ . The curve above the leftward pointing horizontal axis shows the marginal utility for good 2 as a function of its consumption along the IEP (see Appendix A for its derivation). Given $\lambda > 0$ and price p_2 , one locates their product on that horizontal axis, finds the corresponding point on the marginal utility curve, and reads off the vertical axis the optimal consumption M_2 of good 2. Using the IEP in the first quadrant, one then locates M_1 on the rightward pointing horizontal axis. In this example, the budget B encourages the consumer to overspend relative to the lifetime optimal plan.

3.1 Comparative statics

Let us see how consumers using the λ rule react to small changes in the subproblem, and how those reactions compare to those of consumers using the budget rule. During this exercise we hold λ constant, so the elasticities should be regarded as short run.

Income elasticity under the λ rule is trivially zero as long as the consumer is not liquidity constrained. Since she can freely shift purchasing power across the boundary of the subproblem, a marginal shift in income tentatively allocated to the subproblem has no effect on λ and thus no effect on choice. To call out this simple but important observation, we write

Proposition 1. *The (short term) income elasticity of Marshall-Edgeworth demand $M(p, \lambda)$ is zero when the consumer faces no liquidity constraints.*

Now consider price effects. In the case $n = 1$, we have (in current notation) $M(p_1, \lambda) = (u_1)^{-1}(\lambda p_1)$, or using Marshall's enduring convention of putting price on the vertical axis and quantity on the horizontal axis, the demand curve is simply $p_1 = \frac{u_1(x_1)}{\lambda}$. As Marshall explained long ago, one increases consumption of the good until the marginal utility falls to the price scaled by the marginal utility of money.

The same intuition applies to the general case, except that one increases consumption of baskets along the IEP. The impact of a change in a single price is complicated by the implied move to a different IEP. The general case $n > 1$ is a bit awkward to state, but is nicely illustrated in the following result for $n = 2$.

Proposition 2. *Let $n = 2$. Then the sensitivity of the demand function $M(p, \lambda)$ to the price of good 1 is given by*

$$\frac{dM_1(p, \lambda)}{dp_1} = \frac{\lambda u_{22}}{u_{11}u_{22} - u_{12}u_{21}}, \quad (10)$$

$$\frac{dM_2(p, \lambda)}{dp_1} = \frac{-\lambda u_{21}}{u_{11}u_{22} - u_{12}u_{21}}, \quad (11)$$

where all the second partials of u are evaluated at the vector $x = M(p, \lambda)$.

Proof. Recall from (9) that

$$u_1(M(p, \lambda)) = \lambda p_1$$

$$u_2(M(p, \lambda)) = \lambda p_2.$$

Differentiating both sides of both equations with respect to p_1 and solving the resulting system of linear equations we obtain the result. Q.E.D.

Note that the denominator in (10) and (11) is the Hessian determinant, which is positive by the second-order condition for optimality. Of course, the first numerator is negative since $u_{22} < 0$ by concavity, and thus the own price effect is always negative (ruling out Giffen goods). The sign of the cross price effect is the sign of the numerator in (11), determined by the cross partial derivative of u at the optimal consumption. If the goods are substitutes, then $u_{21} < 0$ and a rise in the other good's price will increase the demand of this good, but if they are complements then $u_{21} > 0$ and the same rise will decrease demand. Thus, price effects in $M(p, \lambda)$ arise naturally from the curvature of u , and are easy to interpret and explain.

The case $u_{21} = 0$ of separable goods illustrates the contrast with the budget rule. Begin at an interior point selected by both rules, so $p_2 u_1(x_1) = p_1 u_2\left(\frac{B - x_1 p_1}{p_2}\right)$ from (6), and $\frac{u_1(x_1)}{p_1} = \frac{u_2(x_2)}{p_2} = \lambda$ from (9). Now consider a change in the price of good 1. Inspection of the first expression shows that the budget constrained consumer would have to adjust the consumption of both goods, while the second expression shows that a consumer using the λ rule would only adjust x_1 .

Appendix B details how the impact of a price change can be decomposed into a substitution and an expenditure effect, and how these relate to the standard Hicks-Slutsky decomposition.

3.2 A parametric example

A simple constant elasticity of substitution example illustrates the new Marshallian approach, and suggests why it is far more robust to price surprises than the textbook approach.

Suppose that lifetime preferences are given by $U(x_1, x_2) = \sqrt{x_1} + \gamma\sqrt{x_2}$ with $\gamma \gg 1$. The subproblem is to choose x_1 . This is clearly separable, and the factor γ ensures that it is indeed "small" relative to the (very simplified) rest-of-life problem of choosing x_2 . Suppose that anticipated prices are p_1 and $p_2 \gg p_1$, but $p_2 \leq \gamma p_1$ again consistent with keeping the subproblem small.

Letting $A = \frac{p_2}{\gamma p_1} \leq 1$, routine calculations yield the following formulas for the lifetime optimum:

$$\begin{aligned} x_1^* &= \frac{L}{p_1 + p_2 A^{-2}}, \quad x_2^* = \frac{L}{p_1 A^2 + p_2}, \\ V_o(L, P) &= \sqrt{\frac{L}{p_1 + p_2 A^{-2}}} + \gamma \sqrt{\frac{L}{p_1 A^2 + p_2}}, \\ e &= B = p_1 x_1^* = \frac{L p_1}{p_1 + p_2 A^{-2}}. \end{aligned}$$

Let us derive the marginal utility of money at the optimal expenditure, μ , and our approximation of it, λ : $V(L) = \gamma \sqrt{\frac{L}{p_2}}$, and hence

$$\begin{aligned} \mu &= V'(L - e) = \frac{1}{2p_1} \sqrt{\frac{p_1 + p_2 A^{-2}}{L}}, \\ \text{while } \lambda &= V'(L) = \frac{\gamma}{2\sqrt{L p_2}}. \end{aligned}$$

We shall confirm that the approximation error resulting from evaluating the derivative of the (indirect) utility function V at the original lifetime liquidity is indeed small. Taking the ratio of μ and λ , we have

$$\frac{\mu}{\lambda} = \sqrt{1 + \frac{p_1 A^2}{p_2}} \in \left(1, 1 + \frac{p_1}{2p_2}\right), \quad (12)$$

where we have used $A \leq 1$ and the fact that $\sqrt{1+x} \leq 1 + \frac{x}{2}$, showing that the approximation is close indeed. For example, when $p_2 = 100p_1$, $\frac{\mu}{\lambda} < 1.005$.

Next we calculate the elasticities of the optimal expenditure e and the corresponding marginal utility of money μ relative to the price of today's good p_1 . Straightforward differentiation and some manipulation yields that

$$\frac{de}{dp_1} \cdot \frac{p_1}{e} = -\frac{p_2}{p_1 A^2 + p_2} < -\frac{p_2}{p_1 + p_2} \approx -1. \quad (13)$$

That is, the optimal expenditure is unit elastic, making it rather sensitive to price changes. A 50% increase in today's price will reduce the optimal expenditure for today by roughly 33%, making the budget a poor estimate. On the other hand,

$$\frac{d\mu}{dp_1} \cdot \frac{p_1}{\mu} = -\frac{p_1}{p_1 + p_2 A^{-2}} > -\frac{p_1}{p_1 + p_2} \gg -1. \quad (14)$$

Returning to the example $p_2 = 100p_1$, the elasticity is less than 0.01 (in absolute terms). We can see that the shadow utility of money is highly inelastic, so as a result of a price surprise μ hardly changes and thus it is still well approximated by λ – which of course remains constant as it is determined exclusively by the rest-of-life problem.

Finally it may be worth pointing out that as λ is an underestimate of μ – because V is concave and we use the Taylor approximation below L – it is a marginally better estimate for price increases, that yield a decrease in the marginal utility of money, than price decreases.

4 Using λ

Several practical complications arise when choosing consumption via the marginal utility of money. We will now deal with what we see as the most important complications — indivisible goods, where a marginal analysis does not directly apply; significant shocks to lifetime income, including the purchase of big ticket items; using price observations to adapt λ ; and liquidity constraints.

4.1 Indivisibles

How does the λ rule work when goods are indivisible? Consider first the simplest case: the consumer faces the separable subproblem of whether or not to buy a single indivisible good (or basket of goods) at price p . Indivisibility is captured in the constraint $x \in \{1, 0\}$, and we normalize $u(0) = 0$. Thus the objective function (8) becomes

$$\max_{x \in \{1, 0\}} [u(x) - \lambda xp] = \max \{0, u(1) - \lambda p\}, \quad (15)$$

since $u(x) - \lambda xp = 0 - \lambda 0 = 0$ for $x = 0$. Using the last expression, one can say that the consumer calculates the ratio $\frac{u(1)}{p}$ of perceived quality to price and compares it to λ . If the quality-price ratio, interpreted as value for money, exceeds the marginal utility of money, then she will buy, and otherwise not buy.²

²John Hauser and Glen Urban (1986) pose as alternative hypotheses that consumers use value for money, $\frac{u}{p}$, or “net value,” $u - \mu p$, to prioritize purchases of indivisibles. Our analysis shows that the two rankings are equivalent for yes/no decisions, but we now argue that “net value” is the appropriate ranking expression, as long as $\mu = \lambda$ is properly calibrated.

When the consumer has to choose one out of several mutually exclusive varieties or baskets, the quality-price ratio is no longer sufficient to rank them. A very small basket may offer a high value for money, but still provide only a small utility gain. Instead, the consumer ranks baskets $b^k = (x_1^k, \dots, x_n^k)$ of indivisibles (so each $x_i^k = 0$ or 1) at price vector p according to their net utility gain, $g^k = u(b^k) - \lambda p b^k$. She picks the basket with highest g^k as long as it is positive, and otherwise picks the null basket $b^0 = (0, \dots, 0)$ at price $p b^0 = 0$ and gain $g^0 = u(0) - \lambda p b^0 = 0$. Ignoring trivialities where a more valued basket has lower price, it follows that basket k will be preferred to basket j if and only if $\frac{u(b^k) - u(b^j)}{p b^k - p b^j} \geq \lambda$. The gain in utility by choosing k over j must exceed the shadow utility of the additional expenditure, i.e., the *incremental* quality-price ratio must exceed the marginal utility of money.

By contrast, the budget constrained consumer would pick the highest quality item that does not break her budget. To illustrate the difference, consider the following two scenarios. In the first, the consumer has two baskets available, with $u(b^1) < u(b^2)$ and $p b^1 = B < p b^2$ while $0 < g^2 = u(b^2) - \lambda p b^2 < u(b^1) - \lambda p b^1 = g^1$, so that both decision rules lead to the purchase of basket 1. The second scenario is the same, except that the price of b^2 drops so that $p' b^2 = p b^1 + \varepsilon$. According to the budget rule, basket 2 is still not affordable, so the consumer still buys 1. However, when ε is sufficiently small that $g'^2 = u(b^2) - \lambda p' b^2 > u(b^1) - \lambda p b^1 = g^1$, the consumer using the λ rule would switch to the suddenly relatively inexpensive basket. Crucially, the λ rule would coincide with the new lifetime optimum, while the budget rule would not.

The appearance of a new variety (or a change in the valuation of an existing one) creates a similar contrast. Assume that the perceived quality of the expensive option in the first scenario jumps from $u(b^2)$ to a much higher value, with its price remaining constant. This change would not affect the choice according to the budget rule, but it would clearly lead to – the lifetime optimal – change in behavior according to the λ rule.

4.2 Adjusting λ

The consumer treats λ as a constant in all subproblems considered so far, but there are several situations that require an update, even if a full reoptimization of the life-problem is not necessary.

Let us first consider shocks to lifetime income, which are too large to be ignored but not large enough to necessitate a full reoptimization. In this case the solution is to improve on the precision of the approximation. The second-order Taylor expansion of the indirect utility function tells us what is needed. Suppose that lifetime income “jumps” from L_0 to L_1 . The Taylor approximation around the original lifetime income is $V(y) \approx V(L_0) + (y - L_0)V'(L_0) + \frac{(y - L_0)^2}{2}V''(L_0)$. We can also approximate $V(y)$ by a first-order Taylor expansion around the new lifetime income: $V(y) \approx V(L_1) + (y - L_1)V'(L_1)$. Differentiating both approximations with respect to y , we have $V'(L_0) + (y - L_0)V''(L_0) \approx V'(L_1)$. Finally, letting $y = L_1$ we obtain

$$\lambda_1 = V'(L_1) \approx V'(L_0) + (L_1 - L_0)V''(L_0) = \lambda - \beta\Delta L, \quad (16)$$

where $\Delta L = L_1 - L_0$ is the change in lifetime income and $\beta = -V''(L) \geq 0$ is the rate at which marginal utility diminishes.

Let us quantify the lifetime income effects that we ignore for small expenditures:

Proposition 3. *For two goods, the sensitivity of consumption to lifetime income is given by*

$$\frac{dM_1(p, \lambda)}{dL} \approx -\beta \frac{p_1 u_{22} - p_2 u_{12}}{u_{11} u_{22} - u_{12} u_{21}}, \quad (17)$$

$$\frac{dM_2(p, \lambda)}{dL} \approx -\beta \frac{p_2 u_{11} - p_1 u_{21}}{u_{11} u_{22} - u_{12} u_{21}}, \quad (18)$$

where all the second partials of u are evaluated at the vector $x = M(p, \lambda)$.

Of course β is close to zero when V is approximately linear, so these effects are indeed negligible for small changes of lifetime income. Also note that, just as in the standard model, it is possible that when her lifetime income increases the consumer buys less of one of the goods. This is the case of a backward bending IEP (note that the ratio of (17) and (18) gives us the slope of the IEP), and thus these goods are the same old inferior goods of Marshall.

Proof. We first calculate how the Marshall-Edgeworth demand varies around the optimal choice as λ changes. Recall from (9) that

$$\begin{aligned} u_1(M(p, \lambda)) &= \lambda p_1 \\ u_2(M(p, \lambda)) &= \lambda p_2. \end{aligned}$$

Differentiating both sides of both equations with respect to λ and solving the resulting system of linear equations we obtain

$$\begin{aligned}\frac{dM_1(p, \lambda)}{d\lambda} &= \frac{p_1 u_{22} - p_2 u_{12}}{u_{11} u_{22} - u_{12} u_{21}}, \\ \frac{dM_2(p, \lambda)}{d\lambda} &= \frac{p_2 u_{11} - p_1 u_{21}}{u_{11} u_{22} - u_{12} u_{21}}.\end{aligned}$$

Now note that $\frac{dM_i(p, \lambda)}{dL} = \frac{dM_i(p, \lambda)}{d\lambda} \cdot \frac{d\lambda}{dL}$. Finally, from (16) we have that $\frac{d\lambda}{dL} \approx -\beta$ and our proof is complete. Q.E.D.

The purchase of a big ticket item, like an automobile, has a similar effect to a moderate shock to lifetime income. We can also use a second-order approximation to calculate λ , with the wealth shock substituted by the representative price of the cars under consideration. In case the price range considered is so large that the use of the same λ for all varieties leads to an inaccurate estimate, we can have an individual second order estimate of λ for each potential item, $\lambda_i = \lambda + \beta p_i$. The consumer will then choose the variety, b^i , which maximizes her surplus utility $u(b^i) - \lambda_i p_i$.

4.3 Learning the value of money

As the consumer observes more and more prices, she needs to consider updating her λ based on the new information. We propose a two-step adaptive updating process for λ . The first step is to translate a price observation into news about the value of λ , and the second is to determine the magnitude of the update.

Translation is straightforward for indivisible goods. The quality-price ratio $\frac{u(b^k)}{p b^k}$ is a natural candidate as a new observation of λ . For divisible goods, we need a somewhat different procedure – we would never get a new observation, since x is chosen to satisfy $\frac{u_1(x)}{p_1} = \lambda$. Instead, we recall the quantity chosen “last time”, x^{old} , and evaluate the marginal quality-price ratio at that choice, using the new prices. That is, the new observation of λ is $\frac{u_1(x^{old})}{p_1^{new}}$.

As for the second step, the logical procedure is to “weight” the new observation according to its share in overall consumption, the same way as official inflation figures are calculated.³

³Here we assume, for simplicity, that the consumer does not try to extrapolate from individual observed

In line with the idea that the consumer treats λ as a constant, she will only update it periodically – say, monthly – in possession of m new observations. Define $q_i \in [0, 1]$ as the share of overall expenditure spent on good i in the past “month”. Then the formula to update λ is

$$\lambda' = \left(1 - \sum_{i=1}^m \alpha_i q_i\right) \lambda + \sum_{i=1}^m \alpha_i q_i \lambda_i. \quad (19)$$

Here $\alpha_i \in [0, 1]$ is the parameter measuring how much the consumer weighs new information relative to old. It also captures the consumer’s perception of the permanence of any price changes. Thus a one-off “fire sale” should not carry weight (very low α_i) while a price hike due to a specific tax levied on a product should have an α_i close to one.

Other than the permanence of price change issue, we would expect α_i not to vary with categories but to be larger for individuals whose marginal utility diminishes more quickly.

It may be worth noting that when the consumer buys an indivisible good, the new λ observation will always be above the updated value, while when she does not buy it will always be below. Thus the decision whether to consume will be the same with the updated λ as with the old value.

Note that the updating rule also implies that observations of prices of goods that the consumer does not usually purchase (low q_i) do not affect her view of λ . Similarly, if a good gets priced out of a consumer’s reach, she will stop buying it and this will lead to its exclusion from affecting her view of λ .

4.4 Saving and borrowing

The budget actually plays two distinct roles that often are conflated in textbook analysis. So far we have discussed the budget as targeted expenditure (optimal or otherwise) in the subproblem. The budget also can serve to represent liquidity, a constraint on the purchasing power available in the subproblem, beyond the lifetime constraint $PX = 0$ discussed earlier. This second role can be important even for a new Marshallian consumer, as we will show in this section.

To deal with liquidity issues, we impose a discrete time structure on lifetime consumption prices to changes in the price level. See Angus Deaton (1977) for an exploration of that idea.

and take the subproblem as single period consumption choice. Thus, for time periods $t = 1, 2, \dots$, let the consumer choose consumption bundle $x^t \in \mathfrak{R}_+^n$ and receive a net inflow $x_0^t \in \mathfrak{R}$ of liquid net income. She earns interest at rate $q \geq -1$ on unspent liquid balances, and pays interest at rate $r \geq \max\{0, q\}$ on expenditures in excess of the stock of available liquidity, L_t .⁴ Thus the stock of liquidity evolves according to

$$L_{t+1}(px^t) = L_t + x_0^{t+1} - e(px^t, L_t), \text{ with } e(px^t, L_t) = px^t + r [px^t - L_t]^+ - q [L_t - px^t]^+$$

where $[y]^+ = \max\{0, y\}$. That is, tomorrow's liquidity will be today's liquidity plus tomorrow's liquid income minus today's net expenditure, $e(px, L_t)$, evaluated tomorrow, including interest payments.

The indirect utility (or value) function V in (3) now needs to track liquid assets more carefully. We rewrite it as follows.

$$\widehat{V}_t(L_t) = \max_{x \geq 0} [u(x) + \widehat{V}_{t+1}(L_{t+1}(px))]. \quad (20)$$

As a result of the change in the set of feasible consumption paths, the slope $\widehat{\lambda} = \widehat{V}'_{t+1}(L_{t+1}(px))$, and thus the marginal utility of money, will be different from $\lambda = V'_{t+1}(L)$ in the frictionless case.

Linearizing as before, we have

$$\max_{x \geq 0} [u(x) - \widehat{\lambda}e(px, L_t)], \quad (21)$$

where

$$\widehat{\lambda} = \begin{cases} \lambda(1 + q), & \text{if } L_t > px \\ \lambda(1 + r), & \text{if } L_t < px. \end{cases} \quad (22)$$

The first-order conditions are

$$u_i(x^*) = \widehat{\lambda}p_i, \quad i = 1, \dots, n. \quad (23)$$

The case $r = q$ is essentially the same as before, after rescaling. When $r < q$, however,

⁴We assume constant one-period interest rates, but the reasoning here could be extended to deal with multi-period interest rates that rise with indebtedness and shift downward when additional collateral (illiquid assets) are available. This could lead to a strictly increasing λ as opposed to the simple step function of Figure 3.

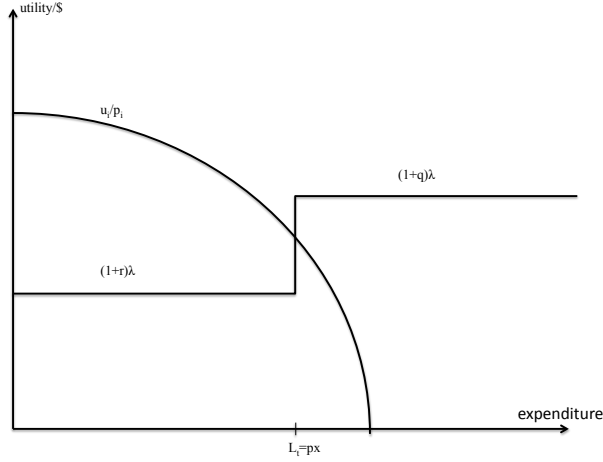


Figure 3: Optimal consumption with liquidity constraints. Expenditure is along the IEP, where u_i/p_i is the same for all goods $i = 1, \dots, n$. In the case shown, optimal expenditure px^* equals available liquidity L_t .

we have in effect a non-constant λ . Figure 3 shows what then happens in the omitted case $L_t = px$ in (22).

A policy implication is that it is better to raise (income) taxes in boom times as they will lead to lower (expected) income effects, since they are less likely to move taxpayers from the saving to the borrowing region of their liquidity.

It is also worth noting that we can no longer claim that a price change has no short term income effects once λ depends on expenditure. Of course, for our step function, we still have that the point elasticity of income is zero almost everywhere, but we can have a better insight by looking at *arc* elasticities. As we prove in Appendix C, the arc elasticities will be still lower under the λ rule than under a budget constraint.

Two extreme cases are of special interest. A paycheck-to-paycheck consumer is unable to borrow and she cannot earn interest on any small saving she might have, so $q = 0$ and $r = \infty$ in (22). She still has to make several purchasing decisions under a single binding budget constraint. At decisions prior to the last one, she should still use the λ rule, as that allows for the lowest level of aggregation. The only difference with respect to consumers with

access to capital markets is that her λ comes from the solution of the remainder of a the pay period, rather than her rest-of-life problem. Current expenditure is no longer insignificant, so λ can move a bit, possibly requiring a second-order approximation.⁵

The other extreme case is the budget-constrained consumer featured in textbooks. He is supposed to be unable to even store any of today's budget $B = L_t$, so $r = \infty$ and $q = -1$ in (22). In Figure 3, the bottom step is on the horizontal axis and the other step is in the sky, so the consumer always chooses to spend exactly $B = L_t$.

5 Applications

5.1 Heuristics

Our discussion so far has been in an orthodox neoclassical vein, but we believe that the new Marshallian approach also contributes to behavioral economics. Even if the lifetime choice problem is never solved and the marginal utility of money is imperfectly calibrated, the λ rule is nevertheless a very convenient and effective rule of thumb.

Convenience is obvious. We have already seen that a single mental yardstick, λ , can be used to choose whether and how much to buy of any isolated good, or of any separable bundle of related goods, whether divisible or indivisible. The same mental yardstick also applies to work decisions. For example, a piece rate worker (or consultant) optimally works until the marginal disutility equals the wage times λ . Likewise the “big ticket” version of the λ rule tells a worker whether to accept a salaried job, or one with fixed hours.

Effectiveness deserves further discussion. Recall that applying fixed budgets to a set of separable subproblems will typically result in a range of shadow prices μ of expenditure. A standard argument shows that the consumer will gain by reallocating expenditure from low

⁵Even so, the λ rule remains superior to the budget rule in the parametric example of Section 3.2. To see this, set $2p_1 = p_2$ in (13) and (14), signifying that today's expenditure is roughly half as much as the remainder until the next paycheck arrives. This would still leave the elasticity of expenditure at $-2/3$, while the elasticity of the marginal value of money would be $-1/3$. The quality of the approximation would also deteriorate $-\frac{\mu}{\lambda} \in [1, 1.25]$ – but not enough to compensate for the difference. Only if there is no need for disaggregation does our method lose its edge.

μ subproblems to those with high μ . By contrast, there is no such deadweight efficiency loss under the λ rule. As long as the miscalibration of λ is no worse than that of the average μ under the budget rule, the λ rule will yield more efficient decisions.

How good an estimate of λ should we expect in typical consumers? Faced with a new subproblem, the consumer need not try to re-solve her lifetime problem (1) or (3). Instead, it can be inferred from her work plan. The marginal utility of leisure divided by the wage is a simple approximation of λ , and it is precisely correct when labor choice is an interior optimum. Previous consumption decisions, of course, also provide convenient estimates of λ ; a consumer who has taken the trouble to optimize once can subsequently free ride on her own calculations.

Of course, some consumers may have a persistently biased view of their true marginal utility for money. A miser is someone who (perhaps because of an impoverished childhood and stalled adjustment) maintains an unreasonably high λ , relative to typical preferences and to his actual lifetime opportunities. Likewise, a spendthrift maintains an unreasonably low λ . Thus the λ rule can be thought of as a heuristic, which is suboptimal to the extent that λ is badly calibrated.

Recent behavioral economics literature includes an alternative heuristic called narrow bracketing, or mental accounting (e.g., Daniel Read, George Lowenstein and Matthew Rabin, 1999; or Richard Thaler, 1999). In our terminology, the claim is that consumers sometimes treat non-separable subproblems as if they were separable, resulting in inefficient choices, especially when self control is problematic. For example, a consumer may regard the health risk as negligible when considering smoking a single cigarette. The addictive properties of nicotine warrant a broader definition of the subproblem, and the consumer might come to a different conclusion when considering smoking a pack a day for a year, or for a decade.

In our view, this literature underlines the importance of separability in defining subproblems. We agree that treating a non-separable subproblem in isolation will typically lead to suboptimal behavior. Our contribution lies mainly in how to proceed after identifying (up to a reasonable approximation) a separable subproblem.

The mental accounting branch of the literature suggests that people may choose a particular budget for each subproblem, e.g., bring exactly \$2500 to Las Vegas. (The ubiquitous

presence of cash machines in casinos suggests that this self-control device is less than 100% successful!) Our recommendation, of course, is instead to make choices according to λ . This doesn't solve all self-control issues, but it has the virtue of simplicity. The consumer then only needs to carry a single number, while a consumer using mental accounting would have to remember a vector. Additionally, we have a simple procedure for how to adapt λ to persistent price changes, while we know of no such method for budgets, mental or otherwise.

When is each heuristic more commonly used? This seems like a good question to take to the laboratory, especially since payments received in lab experiments are normally quite separate from the rest of the subject's life. One could set up various moderately complicated lab tasks, and in one treatment frame the tasks in terms of budgets and in another treatment frame them in terms of the marginal utility of money. Then one could check which treatment encouraged more efficient choices. A more direct test is simply to provide access to budgeting tools as well as λ -oriented tools, and to see which the subjects prefer to use.⁶

5.2 Disaggregated decisions in firms

Multidivisional firms face a question reminiscent of ours: under what conditions can the firm profitably decentralize real investment decisions? To the extent that the divisions are separable, the new Marshallian approach suggests the following decentralized procedure. Let λ be the parent firm's opportunity cost of funds, typically proxied by the weighted average cost of capital. Then each division should invest in scalable projects up to the point that the marginal return falls to λ , and choose among mutually exclusive projects by maximizing present value calculated using λ , the direct generalization of the g rule in section 4.1. When divisions are not separable, the returns should be estimated in terms of incremental net benefit to the firm as a whole, rather than just to the division. Remarkably, this procedure is essentially the same as that recommended in standard corporate finance textbooks (see, for example, Steven Ross, Randolph Westerfield and Bradford Jordan, 2008, Part Four).

Martin Weitzman (1974) discusses the decentralization of decisions on the production side. He compares the relative advantages of quantity targets vs. transfer prices for a production unit that is subject to frequent shocks to its costs. He finds no clear winner as

⁶We are grateful to Ryan Oprea for suggesting this line of inquiry.

the dead-weight losses depend on the relative slopes of marginal cost and marginal benefit curves (much like the incidence of *ad valorem* taxes).⁷ Weitzman agrees with us when he concludes that prices have a comparative advantage when coordination between units is not a major concern and when there are several units producing substitutes.

5.3 Money illusion

Of course, lifetime utility $U(X)$ depends only on actual work and consumption, and is independent of nominal quantities such as L and P . Indeed, the indirect utility function is homogeneous of degree 0 in (L, P) (see, for example, Varian (1992), p. 102). Hence a proportional change in all prices (and wages) will have no effect on the work/consumption plans or utility of *Homo Economicus*. The same is true in the textbook consumer choice subproblem with the budget B adjusted in the same proportion as the price vector.

However, inflation in the real world is much messier than a simultaneous proportional change. An important empirical regularity is that relative prices become more volatile as the measured rate of inflation rises (see, for example, Daniel Heymann and Axel Leijonhufvud, 1995). Actual people updating λ , therefore, are unlikely to immediately adjust to a change in the price level. According to (19) they will overreact to nominal increases in wages and lag in reacting to observed nominal increases in prices of purchased goods (see, for example, David Genesove and Christopher Mayer, 2001). That is, in the parlance of macroeconomic theory, they will suffer from *money illusion*.

Eventually, however, they should adapt fully to the new price level. For a proportional upward change in all prices (and wages, etc.) from P to cP , with $c > 1$, then λ will shift *downward* proportionately to λ/c .⁸ Money is worth less at a higher price level.

⁷The new Marshallian approach would fit into this framework by interpreting λ as the (constant) marginal benefit from the production of an additional unit. For such a situation Weitzman would declare transfer price, that is λ , as the clear winner.

⁸This can be seen from the general result that derivatives of homogeneous degree k functions are homogeneous degree $k - 1$, or verified directly from the definitions: $\lambda' = \lim_{ch \rightarrow 0} \frac{V(cL - ch) - V(cL)}{ch} = c^{-1} \frac{V(L - h) - V(L)}{h} = c^{-1} \lambda$. The middle equality follows from the degree 0 homogeneity of V .

6 Discussion

The new Marshallian approach to consumer choice centers on λ , the marginal utility of money. It applies to any separable subproblem regarding work, or consumption, or both, and λ can be described as the opportunity cost of expenditure in the continuation game following the subproblem. The λ rule states that expenditure on a good (or on the efficient bundles of goods along an appropriate income expansion path, IEP) should increase until its marginal utility diminishes to λ .

The resulting Marshall-Edgeworth demand functions $M(p, \lambda)$ share some features with their standard and Hicksian counterparts $x^*(p, B)$ and $h(p, u)$ – notably, for a given subproblem price vector p , all three lie on the same IEP – but M has several distinct advantages.

- It is very simple and specific – the single number λ is a sufficient statistic for the hugely complex rest-of-life problem. By contrast, each subproblem requires its own budget B in the standard approach, and short of re-solving the lifetime problem, it is unclear how B might be determined.
- It is robust to changes in subproblem prices. No change in λ is required when p changes. By contrast, the appropriate B depends sensitively on p .
- Its elasticities are a first-order approximation of the true General Equilibrium elasticities, and the approximation is exact in a fully separable subproblem whose expenditure is a negligible part of lifetime income. Standard or Hicksian elasticities are close to GE elasticities in such subproblems only in the special case that they coincide with M 's.
- It is quite intuitive – the consumer tries directly to get her money's worth. We suspect that it is also quite descriptive of actual consumer behavior, but we are not aware of any empirical studies so far that compare the predictive power of M and x^* .

To better understand the new Marshallian approach, recall the traditional distinction between non-pecuniary consumption externalities and pecuniary externalities. The first kind recognizes that how much the consumer values a certain quantity of a good may depend on what else is in her consumption basket. Often this kind of externality extends only to a small set of goods, e.g., a few complements and close substitutes, and then we can internalize this

externality in a tractable separable subproblem of the lifetime problem. The second kind, pecuniary externalities, are pervasive and not so easy to encapsulate. Expenditures are mutually exclusive, and money spent on one good is not available to purchase *any* other good. But λ precisely captures this kind of externality.

By contrast, it is hard to specify the appropriate subproblem to which to apply a budget B in the standard approach. The larger the set of goods, the better internalized is the pecuniary externality, and the closer one gets to the true lifetime GE problem. But at the same time, one loses the tractability of low-dimensional partial equilibrium analysis.

The new approach points up the fact that partial equilibrium analysis routinely conflates the available liquidity L with the expenditure B targeted at a subproblem. Otherwise put, the natural subproblem boundaries need not coincide with the boundaries of binding liquidity constraints. Given access to perfectly liquid financial markets, it is erroneous to specify B other than the lifetime budget constraint. When there are additional liquidity constraints, they are best dealt with as in section 4.4, where the textbook analysis is shown to be an extreme and unrealistic special case.

Some readers may object to our use of cardinal utility functions. But the insistence on ordinal utility for consumer choice seems increasingly anomalous, since cardinal utility has become so pervasive elsewhere in economics. Varian's textbook, for example, uses cardinal functions without apology in earlier chapters on production technology and later chapters on risky choice. Why should the sole exception be consumer choice in the absence of risk?

A key insight from the new Marshallian approach is that budgets should not be applied piecemeal to subproblems, but instead all subproblems should be solved consistently using a single sufficient statistic, the marginal utility of money. The point is that funds are fungible, as has long been recognized in other branches of economics.⁹

One final thought. For several generations, economics instructors have tortured undergraduates with various decompositions of income and substitution effects, and picky distinctions between gross and net substitutes and complements, etc. One benefit of the new

⁹In public finance, for example, it is well known that a uniform income tax is more efficient for raising a given amount of revenue than a collection of specific taxes on individual items. We thank Donald Wittman for pointing out the analogy.

Marshallian approach is that it will sweep away such dross, and restore a measure of common sense to consumer choice theory.

7 Notes on related literature

Following the initial intellectual boost by Friedrich Hayek (1945), the issue of partial aggregation was the subject of a lot of (very technical) study more than half a century ago. The idea was to find (nearly) separable parts of an economic problem so that the externalities across parts can be internalized in a manageable manner. Main contributions include Wassily Leontieff (1947), Robert Strotz (1957, 1959), Terence Gorman (1959, 1968), Herbert Simon and Albert Ando (1961), and Gérard Debreu (1960). The latter offers an ordinal definition of separability, essentially that preferences over $x, y \in \xi^{[n]}$ are independent of X^\perp . He shows under quite general conditions that this entails a cardinal (vNM) utility function over separable subproblems.¹⁰

While we also feel inspired by Hayek, our contribution is more conceptual than this literature and does not delve into the technicalities of separability/aggregation more than strictly necessary to get our ideas across.

Ragner Frisch (1959) introduced the idea of demand elasticity holding the marginal utility of money constant (in a representative agent economy). As his interest was the estimation of market demand elasticities in a multi-sector economy, he did not dwell on the justification of the assumption or its implications for consumer behavior.

James Heckman (1974) was the first to use λ -constant comparative statics in a life-cycle model. Truman Bewley (1977) proposed a consumer theory for the life-time problem using the constancy of marginal utility of money. This macro-biased literature is primarily concerned with consumption smoothing and sidesteps the issue of separating off a manageable subproblem and microfounding consumer behavior.

James C. Cox (1975) provides a more complete derivation of the indirect utility function than we do, though with an emphasis on financial assets.

¹⁰Our thanks to Luciano Andreozzi for bringing this paper to our attention.

Xavier Vives (1987) has a kindred contribution to ours in that he also attempts to bring Marshallian ideas back to the mainstream and use them to defend the partial equilibrium approach. He shows that as the number of goods, n , increases, the income effects in any single market are indeed vanishing at at least a rate of $1/\sqrt{n}$.

Finally, József Sákovics (2011) was the first to present a model of consumer choice where budget constraints are soft. In his case the consumers misperceive prices and hence satisfy the budget constraint for these, so actually they either under or overspend.

8 Appendix

8.1 Appendix A: Marginal utility along the IEP

Let $x_1^*(x_2)$ denote the IEP. We know that

$$p_2 u_1(x_1^*(x_2), x_2) \equiv p_1 u_2(x_1^*(x_2), x_2).$$

Differentiating both sides according to x_2 we obtain

$$p_2 \left(\frac{dx_1^*(x_2)}{dx_2} u_{11} + u_{12} \right) = p_1 \left(\frac{dx_1^*(x_2)}{dx_2} u_{21} + u_{22} \right),$$

which simplifies to

$$\frac{dx_1^*(x_2)}{dx_2} = \frac{p_2 u_{12} - p_1 u_{22}}{p_1 u_{21} - p_2 u_{11}},$$

giving us the slope of the IEP.

Now we can derive the formula for the marginal utility of good 2 along the IEP:

$$\frac{du(x_1^*(x_2), x_2)}{dx_2} = \frac{dx_1^*(x_2)}{dx_2} u_1 + u_2 = \frac{p_2 u_{12} - p_1 u_{22}}{p_1 u_{21} - p_2 u_{11}} u_1 + u_2.$$

8.2 Appendix B: Decomposition of price effects

The textbook price effects are ordinal, shaped by the budget constraint, and involve the Slutsky decomposition into substitution and income effects. The standard Slutsky equation

is

$$\frac{\partial x_i(p, e)}{\partial p_1} = \frac{\partial h_i(p, u(p, e))}{\partial p_1} - x_i(p, e) \frac{\partial x_i(p, e)}{\partial e}, \quad (24)$$

where $x(p, e)$ is the standard demand at prices p and available expenditure e , and $h(p, u)$ denotes the Hicksian (expenditure minimizing) demand at prices p and achieved utility u . The first right hand term in (24) is the substitution effect, measuring the rate of movement along a fixed indifference curve onto a new IEP as a result of a price change. The second term is the income effect, the sign of which depends on the slope of the Engel curve, $\frac{\partial x_i(p, e)}{\partial e}$.

The new Marshallian counterpart of (24) can be obtained by differentiating both sides of the identity $M_i(p, \lambda) \equiv x_i(p, e(p, \lambda))$, where $e(p, \lambda)$ is the optimal expenditure given p and λ .

$$\frac{\partial M_i(p, \lambda)}{\partial p_1} = \frac{\partial x_i(p, e)}{\partial p_1} + \frac{\partial x_i(p, e)}{\partial e} \cdot \frac{\partial e(p, \lambda)}{\partial p_1} \quad (25)$$

Substituting (24) into (25) one obtains

$$\frac{\partial M_i(p, \lambda)}{\partial p_1} = \frac{\partial h_i(p, u(p, e))}{\partial p_1} + \frac{\partial x_i(p, e)}{\partial e} \cdot \left(\frac{\partial e(p, \lambda)}{\partial p_1} - x_i(p, e) \right). \quad (26)$$

Equation (26) shows that the substitution effect is the same in both models. This comes as no surprise, as the substitution effect only depends on the new IEP, which – as we have seen – is model independent. The remaining term, which we shall call the *expenditure effect*, is quite different from its textbook counterpart. Instead of arising from an effort to keep nominal expenditure constant as prices change (as in the textbook analysis), it comes from adjusting expenditure for a given marginal utility of money.

Proposition 4. *The difference between the income and the expenditure effects is*

$$D = - \left[x_1(p, e) + \left(p_1 \frac{\partial M_1(p, \lambda)}{\partial p_1} + p_2 \frac{\partial M_2(p, \lambda)}{\partial p_1} \right) \right] \frac{\partial x_i(p, e)}{\partial e}. \quad (27)$$

Proof. Subtract (26) from (24) to obtain the difference is $-\frac{\partial e(p, \lambda)}{\partial p_1} \cdot \frac{\partial x_i(p, e)}{\partial e}$. By definition $e(p, \lambda) = M_1(p, \lambda)p_1 + M_2(p, \lambda)p_2$, and differentiating that identity wrt p_1 yields

$$\frac{\partial e(p, \lambda)}{\partial p_1} = M_1(p, \lambda) + p_1 \frac{\partial M_1(p, \lambda)}{\partial p_1} + p_2 \frac{\partial M_2(p, \lambda)}{\partial p_1}.$$

The consumption of the first good can be written $M_1(p, \lambda) = x_1(p, e)$, so we obtain the desired expression. Q.E.D.

The interpretation of (27) is the following. When a consumer is constrained by a budget, the sensitivity of her expenditure to p_1 (holding utility constant) is

$$\frac{\partial e(p, u)}{\partial p_1} = h_1(p, u) + p_1 \frac{\partial h_1(p, u)}{\partial p_1} + p_2 \frac{\partial h_2(p, u)}{\partial p_1} = h_1(p, u) = x_1(p, e),$$

where the last two terms add to zero by the standard tangency condition at the optimum: $-\frac{p_1}{p_2} = \frac{\partial h_2(p, u)}{\partial h_1(p, u)}$. In other words, as the consumer must keep spending the same amount, the price change has an income effect proportional to the pre-change consumption of the good. When the consumer can freely move purchasing power across the boundary of the subproblem, the expenditure changes in line with the new demand, taking into account the optimal adjustment of consumption.

The following corollary is immediate:

Corollary 1. *Either own or cross price elasticity of demand is larger in the new Marshallian model than in the standard one if and only if either the good in question is normal and $M_1 > \frac{u_2 u_{21} - u_1 u_{22}}{u_{11} u_{22} - u_{12} u_{21}}$, or the good is inferior and $M_1 < \frac{u_2 u_{21} - u_1 u_{22}}{u_{11} u_{22} - u_{12} u_{21}}$.*

Proof. We use Proposition 2 together with the identity $\lambda p_i = u_i$, to obtain that $D = \left[-M_1 - \frac{u_1 u_{22} - u_2 u_{21}}{u_{11} u_{22} - u_{12} u_{21}} \right] \frac{\partial x_i(p, e)}{\partial e}$. Q.E.D.

8.3 Appendix C: Income elasticity with liquidity constraints

Proposition 5. *For normal goods, the (short term) income arc elasticity of demand is strictly lower when the consumer is not budget constrained, but uses the λ rule.*

Proof. Without loss of generality, let us assume that the income shock is positive, denote it by Δ . In that case, the effect on λ will be non-positive – as by normality all the marginal utility functions will shift down – and it will be negative only if the consumer moves from borrowing to saving.¹¹ If she borrowed before then $\mu(B) > \lambda(1 + r)$, if she is saving now then $\mu(B + \Delta) < \lambda(1 + q)$, implying that $\mu(B) - \mu(B + \Delta) > \lambda(r - q)$. As the changes in marginal utility are proportional to the changes in the Lagrange multipliers (c.f. (6)), the proof is complete. Q.E.D.

¹¹In the case that the consumer spends her liquidity exactly, we can impute her a $\hat{\lambda} \in [\lambda(1 + q), \lambda(1 + r)]$, and the same argument goes through (c.f. Figure 3).

It is worth noting that an implication of Proposition 5 is that by allowing for the smoothing of income shocks across periods our model makes the current consumption decisions less responsive to these shocks, providing a possible a micro-foundation for the Permanent Income Hypothesis in a disaggregated environment.

References

- [1] Bentham, J. (1802) *Theory of legislation*. Edited by E. Dumont, translated from the original French by C.M. Atkinson in 1914. Oxford University Press.
- [2] Bewley, T. (1977) "The Permanent Income Hypothesis: A Theoretical Formulation," *Journal of Economic Theory*, 16, 252-292.
- [3] Cox, J.C. (1975) "Portfolio Choice and Saving in an Optimal Consumption-Leisure Plan," *Review of Economic Studies*, XLII(1), 105-116.
- [4] Deaton, A. (1977) "Involuntary Saving Through Unanticipated Inflation," *American Economic Review*, 67(5), 899-910.
- [5] Debreu G. (1960) "Topological Methods in Cardinal Utility Theory," in Arrow, Karlin and Suppes, editors, *Mathematical Methods in the Social Sciences*.
- [6] Frisch, R. (1959) "A Complete Scheme for Computing All Direct and Cross Demand Elasticities in a Model with Many Sectors," *Econometrica*, 27, 177-196.
- [7] Genesove, D. and Mayer, C. (2001) "Loss Aversion and Seller Behavior: Evidence from the Housing Market," *Quarterly Journal of Economics*, 116(4), 1233-1260.
- [8] Gorman, W.M. (1959) "Separable Utility and Aggregation," *Econometrica*, 27(3), 469-481.
- [9] Gorman, W.M. (1968) "Conditions for Additive Separability," *Econometrica*, 36(3/4), 605-609.
- [10] Hauser, J.R. and Urban, G.L. (1986) "The Value Priority Hypotheses for Consumer Budget Plans" *Journal of Consumer Research*, 12(4), 446-462.

- [11] Hayek, F.A. (1945) "The Use of Knowledge in Society," *American Economic Review*, 35(4), 519–530.
- [12] Heckman, J. (1974) "Life Cycle Consumption and Labor Supply: An Explanation of the Relationship between Income and Consumption Over the Life Cycle," *American Economic Review*, 64(1), 88-94.
- [13] Heymann, D. and Leijonhufvud, A. (1995) *High Inflation*, The Arne Ryde Memorial Lectures, Oxford University Press.
- [14] Hicks, J.R. and Allen, R.G.D. (1934) "A Reconsideration of the Theory of Value," *Economica*, Part I: 1(1), 52-76; Part II: 1(2), 196-219.
- [15] Leontieff, W. (1947) "Introduction to a Theory of the Internal Structure of Functional Relationships," *Econometrica*, 15(4), 361-373.
- [16] Marshall, A. (1890/1920) *Principles of economics*. London: Macmillan.
- [17] Read, D., Lowenstein, G. and Rabin, M. (1999) "Choice bracketing," *Journal of Risk and Uncertainty*, 19(1-3),171-197.
- [18] Ross, S.A., Westerfield, R.W and Jordan, B.D., (2008), *Fundamentals of Corporate Finance*, 8th Edition, McGraw-Hill.
- [19] Sákovics, J. (2011) "Reference Distorted Prices," *Quantitative Marketing and Economics*, forthcoming, available on-line.
- [20] Simon, H. A. and Ando, A. (1961) "Aggregation of Variables in Dynamic Systems," *Econometrica*, 29(2), 111-138.
- [21] Stigler, G. J. (1950) "The Development of Utility Theory. I," *Journal of Political Economy*, 58, 307-27.
- [22] Strotz, R.H. (1957) "The Empirical Implications of a Utility Tree," *Econometrica*, 25(2), 269-280.
- [23] Strotz, R.H. (1959) "The Utility Tree—A Correction and Further Appraisal," *Econometrica*, 27(3), 482-488.

- [24] Thaler, R.H. (1999) "Mental Accounting Matters," *Journal of Behavioral Decision Making*, 12(3), 183-206.
- [25] Varian, H.R. (1992) *Microeconomic Analysis*, Third Edition, W.W. Norton and Co., New York.
- [26] Vives, X. (1987) "Small Income Effects: A Marshallian Theory of Consumer Surplus and Downward Sloping Demand," *Review of Economic Studies*, 54, 87-103.
- [27] Weitzman, M.L. (1974) "Prices vs. Quantities," *Review of Economic Studies*, 41(4), 477-491.