Habit Formation and Long-Run Utility Functions*

ROBERT A. POLLAK

Department of Economics, University of Pennsylvania, Philadelphia, Pennsylvania 19174

Received March 5, 1973

This paper extends the work on habit formation of Pollak [Habit formation and dynamic demand functions, J. Polit. Econ. (1970)] and provides a critical counterexample to a conjecture of von Weizsäcker [Notes on endogenous changes of tastes, J. Econ. Theory (1971)] concerning the existence of a "long-run utility function." A linear specification of habit formation is applied to a general system of demand functions with linear Engel curves. It is shown that there exists a utility function which rationalizes the long-run demand functions if and only if they are the steady-state solution to a system of short-run demand functions generated by an additive utility function.

1. INTRODUCTION

This paper extends the earlier work on habit formation of Pollak [6] and provides a critical counterexample to a conjecture of von Weizsäcker [9] concerning the existence of "long-run utility functions."

In [6], I proposed a model of consumer behavior based on habit formation, using a specific class of demand functions. I postulated that some of its parameters depend linearly on past consumption and examined the resulting system of short-run demand functions. From these I found the implied system of long-run demand functions. Although they were defined as the steady-state solution of the system of short-run demand functions, I showed that, in each of the cases I considered, the long-run demand functions could be rationalized by a utility function. This "long-run utility function" is of the same general form as the short-run utility function, although its parameters depend on both the parameters of the short-run utility function and those of the habit formation specification.

All of the short-run utility functions used in Pollak [6] are additive and generate systems of demand functions with linear Engel curves. In [7],

* This research was supported in part by grants from the National Science Foundation. My discussion of the evaluation of welfare when tastes are subject to habit formation owes much to Edwin Burmeister and Stephen A. Ross, to whom I am grateful.

272 Copyright © 1976 by Academic Press, Inc.
All rights of reproduction in any form reserved.
I show that the forms used in Pollak [6] are the only ones exhibiting both of these properties.

The class of utility functions which generates demand functions with linear Engel curves is completely characterized by two functions homogeneous of degree one. When the utility function is also additive, these two functions each assume highly specific $n$-parameter forms, where $n$ is the number of goods. The restriction to a class characterized by two $n$-parameter homogeneous functions of a specific form is obviously a severe one and indicates that only a narrow subclass of the utility functions which generate linear Engel curves are additive.

In this paper I apply the linear specification of habit formation used in Pollak [6] to general systems of demand functions with linear Engel curves. As in Pollak [6], I examine the short-run demand functions and the implied system of long-run demand functions. I show that there exists a utility function which rationalizes these long-run demand functions if and only if they are the steady-state solution to a system of short-run demand functions generated by an additive short-run utility function. That is, of the short-run utility functions which yield systems of demand functions with linear Engel curves, the narrow class of additive ones used as examples in Pollak [6] are the only ones which, under linear habit formation, yield long-run demand functions which can be rationalized by utility functions.

It is not surprising that only a narrow class of short-run utility functions generate long-run demand functions which can be rationalized by a utility function. Gorman [2] examines a continuous-time model of habit formation and shows that the implied long-run demand functions satisfy the Slutsky symmetry conditions only in very special cases. Since Gorman works in continuous time, his results are not directly comparable to the discrete time formulations von Weizsäcker and I consider, but one would not expect the discrete and continuous cases to be fundamentally different.

Von Weizsäcker [9] is primarily concerned with the welfare rather than the positive implications of habit formation. However, his entire discussion rests on his claim that the long-run demand functions can be rationalized by a utility function. In this paper, I show that his claim is not valid, except in some very special cases.

Von Weizsäcker begins with a system of short-run demand functions of the form

\[ q_{1t} = h^1(p_{1t}, p_{2t}, \mu_t, q_{1t-1}, q_{2t-1}), \]

\[ q_{2t} = h^2(p_{1t}, p_{2t}, \mu_t, q_{1t-1}, q_{2t-1}), \]

(1.1)

where $p$'s, $q$'s, and $\mu$'s denote prices, quantities, and total expenditure, hereafter referred to as income. Von Weizsäcker places no restrictions on
the form of the short-run demand functions or on the way the previous period's consumption influences them, but he does restrict his analysis by assuming that there are only two goods:

For simplicity, I shall assume that there are only two goods. Also for simplicity, I shall assume that tastes are only influenced by the consumption vector of the last period. Influences from periods before the last one are neglected. Although the mathematics would become more complicated, I presume that the relaxation of these two assumptions would not change the substance of the argument.

Unfortunately, where integrability conditions are concerned, the assumption that there are only two goods does change the substance of the argument. In particular, the Slutsky symmetry conditions are always satisfied in the two-good case, provided the demand functions satisfy the budget constraint and are homogeneous of degree zero in prices and income. Hence, there is no presumption that either an argument or a result which holds in the two-good case will generalize to the \( n \)-good case, where \( n > 2 \). In fact, neither von Weizsäcker's argument concerning the existence of a long-run utility function nor his result generalize; for \( n > 2 \) the long-run demand functions can be rationalized by a utility function only in certain exceptional cases. Viewed from the perspective of von Weizsäcker [9], this paper demonstrates that the principal results of that paper do not generalize beyond the two-good case. It does this by means of what I have called a "critical counterexample." Of course, any counterexample is sufficient to refute a conjecture, but there are many instances in which a counterexample does not go to the heart of the conjecture and hence leaves open the question of the validity of its principal claim. For example, suppose I conjecture that any real-valued function defined on a closed interval has a zero derivative at its maximum, and you offer, as a counterexample, a function with a boundary maximum. I would respond by adding a clause to my conjecture, restricting it to interior maxima. In this case, the counterexample leads to a refinement of the original conjecture; its scope is restricted, but the basic insight expressed by the original conjecture remains intact. By a critical counterexample I mean one which challenges the central contention of the conjecture. Like any counterexample, it demonstrates that the conjecture is incorrect as it stands. But because it challenges the basic perception of the conjecture, it also shows that the original conjecture cannot be "saved" by adding extra clauses to rule out "pathological" cases. This is not because a valid conjecture cannot be obtained by restricting its scope; usually, a true conjecture can be obtained in this way. The difficulty is that the restricted conjecture would be such an attenuated version of the original that it cannot fairly be said to embody its principal contention. The essential
feature of a critical counterexample is that it forces abandonment of the central assertion of the original conjecture.\(^1\)

It might be argued that my results constitute a critical counterexample to the von Weizsäcker conjecture only for the class of utility functions which generate linear Engel curves. It is conceivable that, from the standpoint of the von Weizsäcker conjecture, linear Engel curves or linear habit formation are “pathological” cases and that the conjecture is generally valid when Engel curves and habit formation are nonlinear. I find this implausible, since there is no apparent reason why these cases should be exceptional. They were chosen as the focus of this investigation, not because it appeared more likely that the conjecture would fail in these cases than in others, but because they are relatively tractable.

The plan of this paper is as follows: In Section 2, I summarize the relevant results on additive utility functions and linear Engel curves from Pollak [7]. In Section 3, I describe the specification of habit formation and investigate its implications for general systems of demand functions with linear Engel curves. In Section 4, I state and prove the central theorem of this paper: Suppose that the short-run demand functions have linear Engel curves, that tastes are subject to linear habit formation, and that the number of goods exceeds two; then the long-run demand functions can be rationalized by a utility function if and only if the short-run utility function is additive. In Section 5, I discuss some of the welfare issues raised by von Weizsäcker and argue that, even when the long-run utility function exists, it has no welfare significance. Section 6 is a brief summary.

2. ADDITIVE UTILITY FUNCTIONS AND LINEAR ENGEL CURVES

In this section I summarize the results on additive utility functions and linear Engel curves obtained in Pollak [7]. A more thorough exposition, proofs, and references to the literature may be found there.

**DEFINITION.** Let \( U(Q) \) be a utility function, where \( Q \) denotes the commodity vector \( (q_1, \ldots, q_n) \).\(^2\)

---

1 For a fascinating discussion of the role of counterexamples in mathematics, see Lakatos [4]. The term “critical counterexample” is my own.

2 We consider only utility functions which are defined in a subset, \( R \), of the commodity space and satisfy the following regularity conditions:

(i) the set of all \( Q \in R \) for which \( U(Q) > U \) is strictly convex for all \( U \).

(ii) \( U \) has strictly positive first-order partial derivatives everywhere in \( R \).

(iii) \( U \) has continuous second- and third-order partial derivatives everywhere in \( R \).
If there exist \( n \) functions, \( u^i(q_i) \), and a thrice differentiable function \( F \), \( F' > 0 \), such that

\[
F[U(Q)] = \sum_{k=1}^{n} u^k(q_k),
\]

then we say that the utility function \( U \) is additive.\(^3\)

**Definition.** Let \( \{h^1(P, \mu), ..., h^n(P, \mu)\} \) denote a system of demand functions, where \( P \) denotes the price vector \((p_1, ..., p_n)\) and \( \mu \) denotes total expenditure or "income."\(^4\) If the demand functions are of the form

\[
h^i(P, \mu) = \chi_i(P) + \gamma_i(P) \mu
\]

in some region of the price-income space, we say that the demand functions are *locally linear in income* or, briefly, *linear*. If

\[
h^i(P, \mu) = \gamma_i(P) \mu,
\]

we say that the demand functions exhibit *expenditure proportionality*. Expenditure proportionality is a special case of linearity.

The income-consumption curves are straight lines in a region of the commodity space if and only if the demand functions are linear; they are straight lines radiating from the origin if and only if the demand functions exhibit expenditure proportionality. Houthakker [3] has pointed out that, if the income consumption curves are linear almost everywhere, then they must either go through the origin or, when account is taken of the non-negativity constraints, they must be broken curves with linear segments. A kink occurs when an increase in income causes the individual to consume a good which he had not previously consumed. This suggests that the linearity hypothesis must be applied with caution, but it is not a valid reason for rejecting it.

In [7] I prove the following.

---

\(^3\) We require that additive utility functions satisfy the additional regularity condition \( u''(q_i) < 0 \), where \( u''(q_i) \) denotes the second derivative of \( u'(q_i) \). This assumption of "diminishing marginal utility" is more restrictive than the convexity requirement of footnote 2. If \( U \) is an additive utility function and \( u''(q_i) > 0 \) (\( u''(q_i) = 0 \)), then an increase in income will cause an increase in the consumption of the \( i \)th good and a decrease (no change) in the consumption of every other good. Since we are concerned with situations in which demand functions are linear in income, there is no serious loss of generality in ruling out this perverse case.

\(^4\) We consider only systems of demand functions which can be generated by utility functions satisfying the regularity conditions of footnote 2.
THEOREM. An additive utility function yields demand functions locally linear in income if and only if it is of one of the three following forms:

\[ U(Q) = \sum_{k=1}^{n} a_k \log(q_k - b_k), \quad a_i > 0, \quad (q_i - b_i) > 0, \quad \sum a_k = 1, \quad (2.3) \]

\[ U(Q) = \sum_{k} \alpha_k (\beta_k + \delta_k q_k)^c, \quad (2.4) \]

\[ U(Q) = -\sum_{k=1}^{n} \alpha_k e^{-\beta_k q_k}, \quad \alpha_i > 0, \quad \beta_i > 0. \]

The utility function (2.3) represents the familiar Klein–Rubin–Stone–Geary linear expenditure system; the implied demand functions are given by

\[ h^i(P, \mu) = b_i - \frac{a_i}{p_i} \sum_k b_k p_k + \frac{a_i}{p_i} \mu. \]

Any admissible utility function of the form (2.4) can be written in one of the following three forms:

\[ U(Q) = -\sum_{k=1}^{n} a_k (q_k - b_k)^c, \quad c < 0, \quad a_i > 0, \quad (q_i - b_i) > 0, \quad (2.7) \]

\[ U(Q) = \sum_{k=1}^{n} a_k (q_k - b_k)^c, \quad 0 < c < 1, \quad a_i > 0, \quad (q_i - b_i) > 0, \quad (2.8) \]

\[ U(Q) = -\sum_{k=1}^{n} a_k (b_k - q_k)^c, \quad c > 1, \quad a_i > 0, \quad (b_i - q_i) > 0. \quad (2.9) \]

(2.7) and (2.8) are translations of the members of the “Bergson family” of utility functions, whose indifference maps correspond to the isoquant maps of the CES class of production functions, while (2.9) includes the additive quadratic. The demand functions corresponding to (2.7)–(2.9) are given by

\[ h^i(P, \mu) = b_i - \gamma_i(P) \sum_{k=1}^{n} b_k p_k + \gamma_i(P) \mu, \]

where \( \gamma_i(P) \) is given by

\[ \gamma_i(P) = \frac{(p_i/a_i)^{1/(c-1)}}{\sum_{k=1}^{n} p_k (p_k/a_k)^{1/(c-1)}}. \]

The fixed coefficient form is also discussed in Pollak [7], but it is excluded by the regularity conditions of footnote 2.
The income-consumption curves corresponding to (2.3), (2.7), and (2.8) radiate upward from the point \((b_1, \ldots, b_n)\), and in these cases the \(b\)'s are often interpreted as necessary or subsistence quantities. Such an interpretation is often useful, but it should be taken figuratively rather than literally, especially since negative \(b\)'s are admissible. The income-consumption curve corresponding to (2.9) radiates downward from \((b_1, \ldots, b_n)\), and in this case the \(b\)'s are interpreted as satiation or bliss points.

The utility function (2.5a) can be regarded as a limiting form of (2.4) since
\[
\lim_{c \to \infty} - \sum_k \alpha_k \left(1 + \frac{-\beta_k}{c} q_k\right)^c = - \sum_k \alpha_k e^{-\beta_k q_k}.
\]

It is convenient to rewrite (2.5a) as
\[
U(Q) = -\sum_{k=1}^n a_k e^{(\alpha_k - \beta_k)/a_k}, \quad a_i > 0 \tag{2.5b}
\]
where \(a_k = 1/\beta_k\) and \(b_k = (1/\beta_k) \log \alpha_k \beta_k\). The corresponding demand functions are
\[
h_i(P, \mu) = b_i - a_i \sum_k p_k a_k b_k - \frac{a_i \mu}{\sum_k p_k a_k} - a_i \log p_i + \frac{a_i \sum_k p_k a_k \log p_k}{\sum_k p_k a_k}.
\tag{2.12}
\]

Gorman [1] has shown that, if an individual's demand functions are linear, then his indirect utility function can be written in the form
\[
\psi(P, \mu) = \frac{\mu}{g(P)} - \frac{f(P)}{g(P)}, \tag{2.13}
\]
where \(f(P)\) and \(g(P)\) are homogeneous of degree one. Because the indirect utility function is ordinal, the phrase "can be written" must be interpreted carefully. Formally,

**Theorem (Gorman).** If an individual's demand functions are linear in income and his preferences can be represented by an indirect utility function, \(\theta(P, \mu)\), then there exists a function \(G, G' > 0\), and functions \(f(P)\) and \(g(P)\), homogeneous of degree one, such that
\[
G[\theta(P, \mu)] = \frac{\mu}{g(P)} - \frac{f(P)}{g(P)}.
\tag{2.14}
\]

As Gorman shows, the demand functions corresponding to (2.14) are given by
\[
h_i(P, \mu) = f_i - \frac{fg_i}{g} + \frac{g_i}{g} \mu, \tag{2.15}
\]
where $f_i(P)$ and $g_i(P)$ denote the partial derivatives with respect to $p_t$ of $f$ and $g$, respectively. The "Gorman forms" of the indirect utility functions corresponding to (2.3)-(2.5) are given by

$$g(P) = \pi p_k^{\gamma_k} \quad \text{and} \quad f(P) = \sum b_k p_k, \quad (2.16)$$

$$g(P) = \left[ \sum a_k p_k^\gamma \right]^{1/c} \quad \text{and} \quad f(P) = \sum b_k p_k, \quad (2.17)$$

$$g(P) = \sum a_k p_k \quad \text{and} \quad f(P) = \sum b_k p_k + \left( \sum a_k p_k \right) \log \left( \sum a_k p_k \right) - \sum a_k p_k \log p_k, \quad (2.18)$$

respectively.

3. LINEAR HABIT FORMATION

In this section I describe the specification of linear habit formation and investigate its implications for systems of demand functions locally linear in income. A more detailed exposition of the linear habit formation model applied to a specific class of demand functions, a proof that the implied system of dynamic demand functions is stable, and references to the literature may be found in Pollak [6].

Following Pollak [6], we introduce habit formation into the linear expenditure system (2.3) by postulating that the $b$'s depend linearly on past consumption. More specifically, we assume that the "necessary" or "subsistence" quantity of good $i$ in period $t$ depends linearly on consumption of that good in the previous period:

$$b_{it} = b_i^* + \beta_i q_{it-1}, \quad 0 \leq \beta_i < 1. \quad (3.1)$$

Here $b_i^*$ can be interpreted as a "physiologically necessary" component of $b_{it}$ and $\beta_i q_{it-1}$ as the "psychologically necessary" component.

If all goods are subject to habit formation of the type described by (3.1), then the demand functions are given by

$$h^{it}(P, \mu, Q^{t-1}) = b_i^* - \frac{a_i}{p_i} \sum p_k b_k^* + \frac{a_i}{p_i} \mu + \beta_i q_{it-1} - \frac{a_i}{p_i} \sum p_k \beta_k q_{kt-1}, \quad (3.2)$$

It should be remarked that the function $f(P)$ is not determined uniquely; nothing is changed when we replace $f(P)$ by $f^*(P) = f(P) + \omega g(P)$.

The requirement $\beta_i < 1$ is a stability condition.
where time subscripts on the $p$'s and $\mu$ have been suppressed. These short-run demand functions, like their static counterparts, are locally linear in income. Since the $b$'s are linear in past consumption and since current consumption depends linearly on the $b$'s, present consumption of each good is a linear function of past consumption of all goods. Since the $b$'s are positive, there is a positive relation between past and current consumption of each good and a negative relation between past consumption of a good and current consumption of every other good.

In this paper I shall consider only the habit formation specification (3.1), which implies that consumption in the previous period influences current preferences and demand but that consumption in the more distant past does not. A more general specification in which the $b$'s depend linearly on a geometrically weighted average of past consumption is considered in Pollak [6], and all of the results obtained in this paper can be extended without difficulty to the more general case.8

This specification of "linear habit formation" can be applied to any system of demand functions locally linear in income. We modify the Gorman form of the indirect utility function (2.13) by replacing $f(P)$ by $f(P) + \sum p_k b_k$; the indirect utility function now becomes

$$
\Psi(P, \mu) = \mu - f(P) - \sum p_k b_k .
$$

(3.3)

The corresponding demand functions are given by

$$
h^i(P, \mu) = b_i - \frac{g_i}{g} \sum p_k b_k + f_i - \frac{g_i}{g} + \frac{g_i}{g} \mu
$$

(3.4)

and are linear in income and in the $b$'s. We now proceed as we did in the case of the linear expenditure system, by postulating that the $b$'s depend linearly on past consumption. This yields a system of short-run demand functions of the form

$$
h^it(P^t, \mu^t, Q^{t-1}) = b_i^* - \frac{g_i}{g} \sum p_k b_k^* + f_i - \frac{g_i}{g} f + \frac{g_i}{g} \mu
$$

$$
+ \beta_i q_{it-1} - \frac{g_i}{g} \sum p_k \beta_k q_{kt-1} .
$$

(3.5)

Given the consumption vector of period zero and given prices and

8 McCarthy [5] shows that a more general system of this type is stable even when different goods have different "memory" coefficients, which is a result I was unable to establish in Pollak [6].
income of period one, the short-run demand functions yield a consumption vector for period one. In a "steady state" or "long-run equilibrium" the optimal consumption vector for period one will be identical with the consumption vector of period zero. And, if prices and income remain constant over time, the optimal consumption vector in every subsequent period will also be equal to the consumption vector of period zero.

The long-run equilibrium consumption vector can be found by solving the short-run demand functions (3.5) under the assumption that \( q_{it} = q_{i,t-1} = q_i \) for all \( i \). To save notation, we replace \( g_i / g \) by \( \gamma^i \). We solve (3.5) for \( q_i \) as a function of the \( p \)'s, \( \mu \), and \( \sum p_k \beta_k q_k \):

\[
q_i = \frac{b_i^* + f_i}{1 - \beta_i} + \frac{\gamma^i}{1 - \beta_i} \sigma, \tag{3.6}
\]

where

\[
\sigma = \mu - f - \sum p_k \beta_k^* - \sum p_k \beta_k q_k.
\]

Multiplying (3.6) by \( p_i \) and summing over \( i \) yields

\[
\mu = \sum p_k q_k = \sum p_k \left( \frac{b_k^* + f_k}{1 - \beta_k} \right) + \sigma \sum \frac{p_k \gamma^k}{1 - \beta_k},
\]

so that

\[
\sigma = \frac{\mu - \sum p_k [(b_k^* + f_k)(1 - \beta_k)]}{\sum [p_k \gamma^k(1 - \beta_k)]}.
\]

Substituting for \( \sigma \) in (3.6), we obtain the long-run demand functions. This proves

**Theorem.** Suppose that the short-run demand functions are locally linear in income (3.4),

\[
q_{it} = b_{it} + f_i(P_t) + \frac{g_i(P_t)}{g(P_t)} \left[ \mu_t - f(P_t) - \sum p_k b_{kt} \right],
\]

and \( b_{it} \) is given by the linear habit function (3.1),

\[
b_{it} = b_i^* \cdot \beta_i q_{it-1}.
\]

Then the long-run demand functions are given by

\[
h^i(P, \mu) = B^i(P) - \Gamma^i(P) \sum p_k B^k(P) + \Gamma^i(P) \mu, \tag{3.7}
\]

where

\[
B^i(P) = \frac{b_i^* + f_i(P)}{1 - \beta_i}. \tag{3.8}
\]
In [6] I considered the implications of linear habit formation for the demand functions generated by the additive utility functions (2.3)–(2.5): all of these yield systems of demand functions with linear Engel curves, and they exhaust the class of additive utility functions which do so. I examined the long-run demand functions and exhibited the utility functions which rationalized them.

**Theorem.** The long-run demand functions corresponding to (2.3) can be rationalized by the utility function

$$U(Q) = \sum A_k \log(q_k - B_k), \quad A_i > 0, \quad (q_i - B_i) > 0, \quad \sum A_k = 1$$

(3.10)

where \( A_i \) and \( B_i \) are given by

$$A_i = \frac{a_i/(1 - \beta_i)}{\sum [a_k/(1 - \beta_k)]}, \quad B_i = \frac{b_i^*}{1 - \beta_i^*}.$$  

(3.11)

The long-run demand functions corresponding to (2.7)–(2.9) can be rationalized by the utility functions

$$U(Q) = -\sum A_k (q_k - B_k)^c, \quad c < 0, \quad A_i > 0, \quad (q_i - B_i) > 0,$$

(3.12)

$$U(Q) = \sum A_k (q_k - B_k)^c, \quad 0 < c < 1, \quad A_i > 0, \quad (q_i - B_i) > 0,$$

(3.13)

$$U(Q) = -\sum A_k (B_k - q_k)^c, \quad c > 0, \quad A_i > 0, \quad (B_i - q_i) > 0,$$

(3.14)

respectively, where \( A_i \) and \( B_i \) are given by

$$A_i = \frac{a_i}{(1 - \beta_i)^{1-c}}, \quad B_i = \frac{b_i^*}{1 - \beta_i^*}.$$  

(3.15)

The long-run demand functions corresponding to (2.5) can be rationalized by the utility function

$$U(Q) = -\sum A_k \exp \left[ \frac{B_k - q_k}{A_k} \right],$$

(3.16)

where \( A_i \) and \( B_i \) are given by

$$A_i = \frac{a_i}{1 - \beta_i}, \quad B_i = \frac{b_i^*}{1 - \beta_i^*}.$$  

(3.17)
The reader should convince himself that the long-run demand functions cannot be rationalized by the utility function obtained by replacing \( q_{it} \) by \( q_i \) and \( b_{it} = b_i^* + \beta_t q_{it-1} \) by \( b_i = b_i^* + \beta_i q_i \) in the short-run utility function. For example, in the case of the linear expenditure system, (2.3), this yields the utility function

\[
\sum a_k \log[q_k - (b_k^* + \beta_k q_k)],
\]

but it is easy to verify that maximizing (3.18) subject to the budget constraint does not yield the long-run demand functions implied by (3.7). The difficulty is that the individual treats \( b_{it} \) as a constant and not as a function of \( q_{it} \), so that maximization with respect to the lagged value of \( q_i \) which appears in \( b_{it} \) is inappropriate.9

In [6] I also showed that, provided \( 0 < \beta_t < 1 \), the dynamic demand functions corresponding to (2.3)-(2.5) are locally stable. If \( Q^0 \), the consumption vector of period zero, is given, then the short-run demand functions determine \( Q^1 \) as a function of \( P^1, \mu_1, \) and \( Q^0 \). In the same way, \( Q^2 \) is determined as a function of \( P^2, \mu_2, \) and \( Q^1 \), or, more conveniently, as a function of \( Q^0, P^1, \mu_1, P^2, \mu_2 \). Thus, for any initial consumption vector \( Q^0 \) and any price-income sequence \( \{(P^1, \mu_1), (P^2, \mu_2), \ldots\} \) the short-run demand functions determine the corresponding consumption sequence, \( \{Q^1, Q^2, \ldots\} \).

The long-run demand functions identify the steady-state consumption vector, \( Q^* \), corresponding to the price-income situation \( (P^*, \mu^*) \). Clearly, if \( Q^0 = Q^* \) and \( \{(P^1, \mu_1), (P^2, \mu_2), \ldots\} = \{(P^*, \mu^*), (P^*, \mu^*), \ldots\} \), then \( \{Q^1, Q^2, \ldots\} = \{Q^*, Q^*, \ldots\} \). In [6], I show that, if \( Q^0 \) is sufficiently close to \( Q^* \), then the consumption sequence corresponding to \( \{(P^*, \mu^*), (P^*, \mu^*), \ldots\} \) will converge to \( Q^* \). The stability proof given in Pollak [6] applies to any system of demand functions locally linear in income; hence, provided \( 0 < \beta_t < 1 \), the stability of the dynamic demand functions (3.4) is guaranteed.

4. THE EXISTENCE OF THE LONG-RUN UTILITY FUNCTION

In this section I prove the central theorem of this paper. Informally, it can be stated as follows: Consider a world with more than two goods, in which the demand functions are locally linear in income and are subject to linear habit formation. In such a world, the long-run demand functions

9 A good illustration is provided by the short-run utility function \( U(Q_t; Q_{t-1}) = V(Q_t) + V^*(Q_{t-1}) \). The long-run utility function is given by \( W(Q) = V(Q) \), not by \( W(Q) = V(Q) + V^*(Q) \).
can be rationalized by a utility function if and only if the short-run demand functions are generated by an additive utility function. More formally,

**Theorem.** Suppose \( n \geq 3 \), that the short-run demand functions are locally linear in income,

\[
q_{it} = b_{it} + f_t(P_t) + \frac{g_t(P_t)}{g(P_t)} \left[ \mu_t - f(P_t) - \sum b_{kt} p_{kt} \right],
\]

where \( f \) and \( g \) are functions homogeneous of degree one with \( g_t(P_t) \neq 0 \) and that \( b_{it} \) is given by the linear habit function (3.1),

\[
b_{it} = b_{i}^* + \beta_t q_{it-1}^{10}.
\]

Then the corresponding long-run demand functions (3.7) can be rationalized by a long-run utility function if and only if

\[
g(P) = \pi p_k^{a_k} \text{ and } f(P) = \sum b_k p_k \tag{2.16}
\]

or

\[
g(P) = \left[ \sum a_k p_k \right]^{1/a} \text{ and } f(P) = \sum b_k p_k \tag{2.17}
\]

or

\[
g(P) = \sum a_k p_k \text{ and } f(P) = \sum b_k p_k + \left( \sum a_k p_k \right) \log \left( \sum a_k p_k \right) - \sum a_k p_k \log p_k. \tag{2.18}
\]

It might appear paradoxical that in each period the short-run demand functions can be rationalized by a short-run utility function, while the long-run demand functions cannot be rationalized by a long-run utility function. Indeed, one might reason that, as the short-run demand functions converge to the long-run demand function, the short-run utility function should converge to the long-run utility function. But this is incorrect. Let \( Q^t = h(P^t, \mu_t, Q^{t-1}) \) denote the short-run demand functions. Given prices \( P^* \) and income \( \mu^* \), the quantity demanded will converge to \( Q^* \) where \( Q^* = h(P^*, \mu^*, Q^*) \). The short-run demand functions are generated by maximizing the short-run utility function \( U(Q^t, Q^{t-1}) \) with respect to \( Q^t \),

\[10\] Ruling out \( g_t = 0 \) excludes goods which are on the border line between inferiority and noninferiority. This avoids problems about division by zero in an already long and messy proof. A similar assumption was made in conjunction with additivity in footnote 3.
but, as we saw in the example at the end of Section 3, the long-run demand functions are not generated by maximizing $V(Q) = U(Q, Q)$ with respect to $Q$. In the limit, the sequence of short-run utility functions approaches a limit—namely, the short-run utility function $U(Q, Q^*)$—and maximizing this short-run utility function with respect to $Q$ generates the short-run demand functions $Q = h(P, \mu, Q^*)$. At $(P^*, \mu^*)$ these short-run demand functions coincide with the long-run demand functions, but, in general, they coincide only at the point $(P^*, \mu^*)$ and not for all $(P, \mu)$, even in a neighborhood of $(P^*, \mu^*)$.

**Proof.** Necessity follows directly from the results cited in Sections 2 and 3. If the indirect utility function is of the form (2.16), (2.17) or (2.1), then the direct utility function is given by (2.3), (2.4), or (2.5) (Pollak [7]). But each of these cases leads to long-run demand functions which can be rationalized by a utility function (Pollak [6]).

Sufficiency is more difficult. We first show that $g$ is of the required form. If the long-run demand functions (3.7) can be rationalized by a utility function, then they can be written in the Gorman form

$$h^i(P, \mu) = F_i(P, \beta, b^*) - \frac{G_i(P, \beta, b^*)}{G(P, \beta, b^*)} F(P, \beta, b^*) + \frac{G_i(P, \beta, b^*)}{G(P, \beta, b^*)} \mu,$$

(4.1)

since they are locally linear in income. Notice that $F$ and $G$ are functions of $P$, $\beta$, and $b^*$ and are homogeneous of degree one in $P$. Equating (4.1) and (3.7) and differentiating with respect to $\mu$ yields

$$\Gamma^i(P, \beta) = \frac{G_i(P, \beta, b^*)}{G(P, \beta, b^*)}.$$

(4.2)

Calculating $\Gamma^i_j$ and $\Gamma^i_j$ from (4.2), we find that

$$\Gamma^i_j(P, \beta, b^*) = \Gamma^i_j(P, \beta, b^*),$$

(4.3)

where subscripts denote partial derivations with respect to price. This relation must hold as an identity in the $\beta$'s and $b^*$'s as well as the $P$'s. To simplify our notation, let

$$S = \sum \frac{P_k \gamma^k}{1 - \beta_k}.$$

We calculate from (3.9)

$$\Gamma^i_j(P, \beta, b^*) = \left[ \frac{\gamma^i_j}{(1 - \beta_j)} S - \frac{\gamma^i_j}{1 - \beta_i} \sum \frac{P_k \gamma^k_j}{1 - \beta_k} S^{-2} - \frac{\gamma^i_j}{1 - \beta_i} \frac{\gamma^i_j}{1 - \beta_j} S^{-2} \right].$$
The last term is symmetric in $i$ and $j$, so (4.3) implies that the factors in brackets must also be symmetric. Equating the term in brackets with the corresponding term in $T^t$ and differentiating with respect to $\beta_t$, $t \neq i, j$, yields
\[
\frac{1}{1 - \beta_t} [\gamma_i^t \gamma^t - \gamma^t \gamma_i^t] = \frac{1}{1 - \beta_t} [\gamma_j^t \gamma^t - \gamma^t \gamma_j^t].
\]
Differentiating with respect to $\beta_t$ implies
\[
\gamma_i^t \gamma^t - \gamma^t \gamma_i^t = 0.
\]
Replacing $\gamma^t$ by $g_t/g_i$, we find that this implies
\[
g_{ij}g_t - g_{ij}g_t = 0,
\]
but
\[
\frac{\partial}{\partial \beta_t} \left(\frac{g_t}{g_i}\right) = 0
\]
if and only if (4.4) holds. Hence, $g(P)$ is additive in the ordinal sense. But $g(P)$ is also homogeneous of degree one, and the only functions which are both additive and homogeneous of degree one are those of the Bergson family:
\[
g(P) = \pi p_k^a, \quad \sum a_k = 1, \quad (4.5)
g(P) = \left(\sum a_k p_k^c\right)^{1/c}, \quad (4.6a)
g(P) = \sum a_k p_k. \quad (4.6b)
\]
Clearly, (4.6b) is a special case of (4.6a), corresponding to $c = 1$. The Cobb–Douglas case (4.5) is the limiting case corresponding to $c = 0$. Thus, $g$ must be of one of the three forms asserted in the theorem.

Unfortunately, we cannot identity $B^t$ with $F_t$. On reflection, this is as it should be, since the Gorman form does not uniquely identify $F$. Since $f = \tilde{f}$ and $f = \tilde{f} + Dg$, where $D$ is any constant, imply the same demand functions, our theorem’s claims about the form of $f$ must be understood to refer to the canonical form in which $D = 0$.

Let $T^t$ be defined by
\[
T^t = B^t - T^t \sum p_k B^k. \quad (4.7)
\]
If the long-run demand functions are theoretically plausible, then
\[
T^t + \Gamma^t \mu = F_t - \frac{G_i}{G} F + \frac{G_i}{G} \mu.
\]
From (4.2), \( T^i = G_i/G \), so we can subtract the income terms and obtain
\[
T^i = F_i - \frac{G_i}{G} F = F_i - \Gamma^i F.
\]

We calculate
\[
T^i = F_{ji} - \Gamma^i F - \Gamma^j F_i
\]
and
\[
\Gamma^i T^i - T^i = (\Gamma^i F_j + \Gamma^j F_i) + (\Gamma^i F - F_{ji} - \Gamma^i \Gamma^j F).
\]

Equation (4.8) is symmetric in \( i \) and \( j \) since \( \Gamma^i_j = \Gamma^j_i \) from (4.2).

To pursue the implications of the symmetry of (4.8), we replace \( \beta_i \) by \( \alpha_i = 1/(1 - \beta_i) \). Replacing \( T^i \) by (4.7), where the \( B_i \)'s are given by (3.7), replacing \( T^i \) by the corresponding partial derivative of \( T^i \), and dropping terms which are clearly symmetric in \( i \) and \( j \) yield
\[
-\alpha_{ij} f_{ji} + \Gamma^i \sum p_k \alpha_k f_{ki}.
\]

Since (4.9) was obtained from (4.8) by dropping symmetric terms, (4.9) must also be symmetric in \( i \) and \( j \). Replacing \( \Gamma^i_j \) by the corresponding expressions involving \( g \)'s, (4.9) becomes
\[
\alpha_{ij} f_{ji} \sum \frac{\alpha_j g_{ij}}{\sum p_k \alpha_k g_k} - \sum p_k \alpha_k f_{ki},
\]
and the symmetry implies
\[
-\alpha_{ij} f_{ji} \sum p_k \alpha_k g_k + \alpha_j g_{ij} \sum p_k \alpha_k f_{ki} = -\alpha_i f_{ij} \sum p_k \alpha_k g_k + \alpha_i g_{ij} \sum p_k \alpha_k f_{kj}.
\]

Differentiating with respect to \( \alpha_s, s \neq i, j \), and dividing by \( p_s \) yield
\[
-\alpha_{ij} f_{ji} g_s + \alpha_j g_{ij} f_{si} = -\alpha_i f_{ij} g_s + \alpha_i g_{ij} f_{sj}.
\]

Differentiating with respect to \( \alpha_j \) yields
\[
-f_{ij} g_s + g_{ij} f_{si} = 0,
\]
so, changing subscripts in the interest of clarity,
\[
\frac{f_{is}}{f_{ir}} = \frac{g_s}{g_r}, \quad s, r \neq i.
\]

We now use the homogeneity of \( f \) and \( g \) to determine a similar restriction involving \( f_{si} \). From (4.10)
\[
f_{is} g_r = g_s f_{ri}, \quad r, s \neq i.
Multiplying by \( p_s \) and summing over \( s, s \neq i \), yields

\[
g_r \sum_{s \neq i} p_s f_{si} = f_{ri} \sum_{s \neq i} p_s g_s. \tag{4.11}
\]

But since \( g \) is homogeneous of degree one,

\[
\sum_{s \neq i} p_s g_s = g - p_i g_i.
\]

Since \( f_i \) is homogeneous of degree zero,

\[
\sum_{s \neq i} p_s f_{is} = -p_i f_{ii}.
\]

Hence, (4.11) implies

\[
\frac{f_{ii}}{f_{ir}} = -\frac{g - p_i g_i}{p_i g_r}. \tag{4.12}
\]

It is easy to verify that (4.10) and (4.12) imply

\[
f_i = \varphi_i \left( \frac{g}{p_i} \right), \tag{4.13}
\]

where the \( \varphi \)'s are arbitrary functions. We must now establish restrictions on the \( \varphi \)'s. Since \( f_{ij} = f_{ji} \), (13) implies

\[
\psi' \left( \frac{g}{p_i} \right) g_i = \psi' \left( \frac{g}{p_j} \right) g_i. \tag{4.14}
\]

Hence,

\[
\frac{\psi' \left( g | p_i \right)}{p_i g_i} = R(P), \quad i = 1, \ldots, n,
\]

where \( R \) is a function of \( P \), independent of the index \( i \). The left-hand side of (4.14) is homogeneous of degree \(-1\), so \( R \) is a function homogeneous of degree \(-1\). Hence, if \( R \) is a constant, it must be 0; we begin by considering that case.

If \( R(P) = 0 \), then (4.14) implies

\[
\psi' \left( \frac{g}{p_i} \right) = 0,
\]

so \( \psi' \left( g | p_i \right) = b_i \). Hence, \( f_i(P) = b_i \), and, since \( f = \sum p_k f_k \),

\[
f(P) = \sum p_k b_k,
\]

which is consistent with our theorem.
If $R \neq 0$, we differentiate (4.14) with respect to $p_t$, $t \neq i$, and, forming ratios, we find

$$\frac{\Psi^{i\nu'} p_t}{\Psi^{i\nu'}} = \frac{g(R_t g_t + R_{g_t})}{R_t g_t} = \frac{R_t g + R_{g_t}}{R_t g} + \frac{g_{g_t}}{g_t g_t}. \quad (4.15)$$

But, since $g$ is given by (4.5), (4.6a), or (4.6b),

$$\frac{g_{g_t}}{g_t g_t} = 1 - c$$

and is independent of $P$ and the indexes $i$ and $t$. Since the right-hand side of (4.15) is independent of the choice of $t$, we must have

$$\frac{R_t g}{R_t g} = \frac{R_s g}{R_s g} \quad (4.16)$$

for all $s \neq i$.

Since $R$ is not constant and is homogeneous of degree zero, we may pick an $s \neq i$ for which $R_s \neq 0$ and write (4.16) as

$$\frac{R_t g}{R_s g} = \frac{g_t}{g_s},$$

which implies

$$R = \frac{D}{g},$$

where $D$ is a constant. Hence,

$$\frac{R_t g}{R_t g} = -1,$$

and (4.15) becomes

$$\frac{\Psi^{i\nu'}(z_t)}{\Psi^{i\nu'}(z_t)} = -\frac{c}{z_t}. \quad (4.17)$$

Hence,

$$\Psi^{i\nu'}(z_t) = d_i z_t^{-c}. \quad (4.18)$$

We now treat separately the cases of (i) $c \neq 1$ and (ii) $c = 1$.

(i) If $c \neq 1$, (4.18) implies

$$\Psi^{i}(z_t) = b_i + d_i z_t^{1-c},$$
where \( d_i = \overline{a_i}/(1 - c) \). This implies

\[
f_i = b_i + d_i \left( \frac{g_i}{p_i} \right)^{1-c} \tag{4.19}
\]

and

\[
f = \sum p_k b_k + g^{1-c} \sum d_k p_k^c. \tag{4.20}
\]

If \( c = 0 \), (4.20) becomes

\[
f = \sum p_k b_k + Dg,
\]

where \( D = \sum d_k \) and is consistent with our theorem.

If \( c \neq 0 \), differentiating (4.20) with respect to \( p_i \) yields

\[
f_i = b_i + (1 - c) g^{-c} g_i \sum d_k p_k^c + g^{1-c} d_i p_i^{c-1}. \tag{4.20}
\]

Replacing \( f_i \) by (4.19) and subtracting \( b_i \) yields

\[
d_i g^{1-c} p_i^{c-1} = (1 - c) g^{-c} g_i \sum d_k p_k^c + d_i c g^{1-c} p_i^{c-1},
\]

so that

\[
d_i p_i^{c-1} g = g_i \sum d_k p_k^c. \tag{4.21}
\]

There are two subcases to consider: (ia) If \( \sum d_k p_k^c = 0 \) for all \( p \), then \( d_i = 0 \) for all \( i \) and (4.20) implies \( f = p_k b_k \), as asserted in our theorem. (ib) If \( \sum d_k p_k^c \neq 0 \) for all \( p \), then from (4.21), \( d_i \neq 0 \) for any \( i \). Hence, (4.21) implies

\[
\frac{g_i}{g_j} = \frac{d_i p_i^{c-1}}{d_j p_j^{c-1}}.
\]

But from (4.6)

\[
\frac{g_i}{g_j} = \frac{a_i p_i^{c-1}}{a_j p_j^{c-1}},
\]

so

\[
d_i = Da_i, \quad i = 1, ..., n.
\]

Hence,

\[
\sum d_k p_k^c = D \sum a_k p_k^c = dg^c,
\]

and (4.20) becomes

\[
f = \sum b_k p_k + Dg
\]

as asserted by our theorem.
(ii) If $c = 1$, (4.18) implies

$$
\Psi^h(z_i) = b_i + d_i \log z_i ,
$$

where $d_i = \bar{d}_i$. This implies

$$
f_i = b_i + d_i \log g - d_i \log p_i
$$

(4.22)

and

$$
f = \sum b_k p_k + \left( \sum d_k p_k \right) \log g - \sum d_k p_k \log p_k .
$$

(4.23)

Differentiating (4.23) with respect to $p_i$ yields

$$
f_i = b_i + d_i \log g + \frac{g_i}{g} \sum d_k p_k - d_i - d_i \log p_i .
$$

(4.24)

Equating (4.22) and (4.24) and subtracting $b_i + d_i \log g - d_i \log p_i$, we find

$$
\frac{g_i}{g} \sum p_k d_k = d_i .
$$

(4.25)

There are two subcases to consider: (iia) If $\sum d_k p_k = 0$ for all $P$, then $d_i = 0$ for all $i$ and (4.2) implies $f = b_k p_k$. (iib) If $\sum d_k p_k \neq 0$ for all $P$, then, from (4.25), $d_i \neq 0$ for any $i$. Hence, (4.25) implies

$$
\frac{g_i}{g_j} = \frac{d_i}{d_j} ,
$$

which implies that the $d$'s are proportional to the $a$'s:

$$
d_i = Da_i , \quad i = 1, \ldots, n .
$$

Hence,

$$
\sum d_k p_k - D \sum a_k p_k ,
$$

and (4.23) becomes

$$
f = \sum b_k p_k + D \left( \sum a_k p_k \right) \log \left( \sum a_k p_k \right) - D \sum a_k p_k \log p_k .
$$

(4.26)

We must now show that the class of indirect utility functions defined by (4.6b),

$$
g(P) = \sum a_k p_k ,
$$

and (4.26) is equivalent to (2.18), although the $a$'s are defined differently
in (4.6b) and (4.26) than in (2.18). To show the equivalence of the two classes, it is useful to rewrite (4.6b) and (4.26) replacing \(a_i\) by \(\bar{a}_i\); we do this to make it clear that we are asserting the equivalence of the classes of utility functions and not the identity of particular coefficients. Rewriting (4.6b) and (4.26) yields

\[
\begin{align*}
g &= \sum \bar{a}_k p_k, \\
f &= \sum b_k p_k + D \left( \sum \bar{a}_k p_k \right) \log \left( \sum \bar{a}_k p_k \right) - D \sum \bar{a}_k p_k \log p_k. 
\end{align*}
\]

(4.6b')

We next observe that we can multiply \(g\) by a positive constant without altering the indifference map. We define \(a_i\) by

\[
a_i = D\bar{a}_i
\]

and rewrite (4.6b') and (4.26') as

\[
\begin{align*}
g &= \sum a_k p_k, \\
f &= \sum b_k p_k + \left( \sum a_k p_k \right) \log \left( \sum a_k p_k \right) - \sum a_k p_k \log p_k \\
&\quad - \left( \sum a_k p_k \right) \log D. 
\end{align*}
\]

(4.6b'')

(4.26'')

But the Gorman form does not uniquely identify \(f\), and there is no way to distinguish between \(f = \bar{f}\) and \(f = \bar{f} + kg\), where \(k\) is any constant. Since the last term of (4.26'') is \(-g \log D\), (4.6b'') and (4.26'') represent the same preference ordering as (2.18).

5. WELFARE IMPLICATIONS

Von Weizsäcker argues that the long-run utility function is the appropriate criterion by which to judge the welfare effects of changes in consumption. Since we have shown, contrary to his conjecture, that when there are more than two goods the long-run utility function exists only in special cases, it cannot serve as a general welfare criterion. Revealed preference arguments provide no help; the counterpart of the nonexistence of the long-run utility function is violation of the long-run version of the strong axiom of revealed preference. At most, then, the long-run utility function may be the appropriate welfare criterion only for a narrow class of cases in which it exists, and I shall argue that there are compelling objections to it even in these cases.
Von Weizsäcker [9, p. 352] shows that the long-run indifference curves are related to binary choice behavior:

In other words, it is possible for the consumer to go from $Q^a$ to $Q$ in a finite number of periods and always feel improved compared to the already attained status quo of the last period if and only if $Q$ lies above the indifference curve going through $Q^a$.

Let $\pi(Q^*)$ denote the short-run preference relation given that consumption in the previous period was $Q^*$; then $Q^a \pi(Q^*) Q^b$ means that the consumption vector $Q^a$ is preferred to $Q^b$ on the basis of the short-run preference ordering corresponding to past consumption of $Q^*$. Von Weizsäcker's assertion, then, is that there exists a finite sequence of vectors $\{Q^0, Q^1, \ldots, Q^n = \bar{Q}\}$ such that

$$Q^1 \pi(Q^0) Q^0, Q^a \pi(Q^a) Q^1, \ldots, \bar{Q} = Q^a \pi(Q^{n-1}) Q^n$$

if and only if $\bar{Q}$ lies on a higher long-run indifference curve than $Q^0$. He does not claim that (5.1) holds for every sequence of vectors $\{Q^0, Q^1, \ldots, Q^n = \bar{Q}\}$—clearly, it does not. Nor does he claim that it is true for the doubleton sequence $\{Q^1, \bar{Q}\}$. He claims only that there exists at least one sequence satisfying (5.1) if and only if $\bar{Q}$ lies on a higher indifference curve than $Q^0$.

Von Weizsäcker continues:

In this sense, the long-run indifference curves exhibit the "long-run preference structure" of the person.

Presumably this is intended as a definition of the "long-run preference structure." It is not a convincing argument that the long-run utility function coincides with our intuitive notion of an individual's long-run preference ordering. To see why, we must examine the implications of habit formation for binary choice behavior and the relation between binary choice and the long-run utility function.

First consider the short run. Let $\pi(Q^a)$ and $\pi(Q^b)$ denote the short-run preference relations implied by consumption of $Q^a$ and $Q^b$ in the previous period. It is certainly possible to have $Q^a \pi(Q^a) Q^b$ and $Q^b \pi(Q^b) Q^a$. That is, $Q^a$ is preferred to $Q^b$ on the basis of the preference ordering induced by past consumption of $Q^a$, while $Q^b$ is preferred to $Q^a$ on the basis of the preference ordering induced by past consumption of $Q^b$. For example, suppose the short-run utility function is the linear expenditure system

$$U(q_{1t}, q_{2t}) = \frac{1}{2} \log (q_{1t} - b_{1t}) + \frac{1}{2} \log (q_{2t} - b_{2t})$$

where $b_{1t} = \frac{1}{2} q_{1t-1}$. It is easy to verify that, if $Q^a = (6, 4)$ and $Q^b = (4, 7)$,
then \( Q^a \pi(Q^a) Q^b \) and \( Q^b \pi(Q^b) Q^a \). This is not a contradiction; it is the essence of the habit hypothesis. Preferences depend on past consumption, and different consumption histories imply different short-run preference orderings.

In this case the long-run demand functions can be rationalized by the long-run utility function

\[
U(q_1, q_2) = \frac{1}{2} \log q_1 + \frac{1}{2} \log q_2
\]  

or, equivalently,

\[
U^*(q_1, q_2) = q_1 q_2.
\]

Hence, \( Q^b = (4, 7) \) is on a higher long-run indifference curve than \( Q^a = (6, 4) \). We have already seen that in a binary choice situation the individual's preference between (6, 4) and (4, 7) depends on his past consumption. This does not contradict von Weizsäcker's assertion that there exists a sequence of consumption vectors by which the individual can go from \( Q^a \) to \( Q^b \) in a finite number of periods and always feel improved (on the basis of his current short-run tastes) compared to the consumption vector of the previous period. It does imply that von Weizsäcker's assertion is true only for carefully chosen sequences, and that it need not be true for the sequence consisting of only \( Q^a \) and \( Q^b \). However, congruence with binary choice is the sine qua non of a utility function representing preferences, so the dependence of the individual's preference between (6, 4) and (4, 7) on his consumption history casts serious doubt on the possibility of interpreting his choice behavior in terms of a consistent preference ordering.

It might be objected that focusing on a single binary choice between (6, 4) and (4, 7) fails to capture the long-run aspect of the situation. But if the individual is offered a finite or an infinite sequence of binary choices between these two collections, then his choice will still depend on his initial consumption pattern. Suppose he is offered an infinite sequence of binary choices in which, at each decision point, he must choose between (6, 4) and (4, 7), and that his initial consumption pattern was (6, 4); then at each decision point he will choose (6, 4). Similarly, if his initial consumption pattern was (4, 7), then at each decision point he will choose (4, 7).

The issue, then, is whether the existence of a von Weizsäcker sequence which enables us to go from \( Q^b \) to \( Q \) in a finite number of steps, feeling improved at each step, implies that the individual is better off—in terms of his preferences—at \( Q \) than at \( Q^b \). I interpret the individual's willingness to move from \( Q^b \) to \( Q \) in a sequence of small steps when he is unwilling to do so in a single large step as indicative of his failure to understand the habit formation mechanism and not of the underlying superiority of \( Q \).
For example, a nonsmoker might prefer to remain a nonsmoker rather than smoke three packs of cigarettes a day, but he might choose to smoke half a pack a day rather than abstain completely. After becoming accustomed to smoking half a pack a day, the individual might prefer to remain a light smoker rather than smoke three packs a day, but he might choose to smoke a pack a day rather than continue at half a pack a day. By this process, the myopic nonsmoker is led to become a heavy smoker. This scenario is entirely consistent with von Weizsäcker's assumptions, yet I am loath to conclude that the individual is better off at $\bar{Q}$ than at $Q^0$. Note that the issue is not whether cigarette smoking is harmful to his health; according to our individualistic welfare premise, if a man prefers three packs a day to none, we would say that he is better off with three. The example, with appropriate repackaging, applies equally to cigarettes, candy, or artichokes.

To recapitulate, the individual's long-run demand behavior is consistent with the hypothesis that he is maximizing the long-run utility function (5.3), but we cannot conclude that a movement from a lower indifference curve to a higher one—e.g., from $(6, 4)$ to $(4, 7)$—would make him better off. If the individual is offered a nonmarket choice between these two baskets—for example, a political choice in which he must vote for one or the other of these two consumption patterns—then his choice will depend on his past consumption history. If his previous consumption has been $(6, 4)$, then he will prefer $(6, 4)$; if it has been $(4, 7)$, he will prefer $(4, 7)$. The fact that the long-run utility function assigns a higher value to one than the other does not imply that it will be selected in this binary choice situation, and it does not imply that it is a point of higher welfare.

From the standpoint of positive economics, the fact that the long-run demand functions were not generated by maximizing the long-run utility function (5.3) is completely irrelevant; the utility function is just a convenient device for coding all of the information about demand behavior. From the standpoint of normative economics, an individual's utility function is significant because it represents his preferences. If we assume that social welfare depends on individual welfare and that each individual is the sole judge of his own welfare, then individual utility functions are germane to welfare economics. Ordinarily, there is no need to distinguish between these two interpretations of the utility function. If a utility function rationalizes an individual's demand functions, we usually assume that it also represents his preferences. In the habit formation model, this is an unwarranted assumption; there is no long-run preference ordering, and the distinction between the demand and preference interpretations of the utility function is critical.

The long-run utility function is the same type of construct as a com-
munity indifference map which rationalizes market demand functions; if it exists, it is a convenient device for coding all of the information about demand behavior, but this is all (see [8]). In general, market demand functions cannot be rationalized by a "market utility function." In those special cases in which they can be, the utility function must be scrupulously interpreted in terms of positive economics; it has no normative or welfare significance.

Since von Weizsäcker's long-run utility function approach to the evaluation of welfare is technically possible only in a narrow class of cases and is conceptually unsatisfactory even for those, it is desirable to consider alternative approaches. One such alternative, one which von Weizsäcker suggests in his concluding paragraph, is to view the problem in an inter-temporal framework. That is, instead of focusing on the one-period utility function, $U(Q_t, Q_{t-1})$—assumed to be the same in every period—we evaluate welfare in terms of the intertemporal utility function

$$V(Q) = W[U(Q_1, Q_0), U(Q_2, Q_1), \ldots, U(Q_T, Q_{T-1})].$$  

(4.3)

The proposed welfare test, in other words, is whether the individual—taking full account of the impact of his present consumption on his future tastes—would be willing to undertake a particular change.

The principal difficulty with this approach is that it is schizophrenic. The habit formation model is tractable because the individual is assumed to be myopic—he fails to recognize the impact of his current consumption on his future tastes. If he could be persuaded to recognize the effects of habit formation, then it would certainly be appropriate to base welfare comparisons on the intertemporal utility function, but then his demand behavior would be far more complex than that predicted by the model of "myopic" habit formation. If an individual insists on being myopic, it is less clear that the intertemporal utility function is the appropriate welfare criterion, but it is tempting to take a paternalistic view and argue that it is. However, it is difficult to reconcile this approach to welfare with an approach to demand analysis based on myopic habit formation.

A second alternative to von Weizsäcker's approach retains the one-period framework, but, instead of beginning with the short-run utility function and an assumption about the way it depends on past consumption, it begins with a long-run (single-period) utility function and a lagged adjustment hypothesis. The latter is invoked to describe the movement from one long-run equilibrium to another. Only in a narrow class of cases will there exist a utility function which rationalizes the short-run demand functions, but the welfare analysis can be conducted on the basis of the long-run utility function. I find the assumption that an individual's
short-run behavior is consistent with a preference ordering more plausible than the corresponding assumption about long-run behavior, but this is clearly an empirical matter. The direct specification of a long-run utility function and an adjustment hypothesis seems to correspond more closely to von Weizsäcker's verbal exposition in his Section 5 than do the short-run utility-function-habit-formation models which provide the basis for his mathematical analysis and mine.

6. Summary

The two principal contentions of this paper are: first, in a model of habit formation with more than two goods, only in a narrow class of cases does there exist a utility function which rationalizes the long-run demand functions. This contention is supported by a critical counterexample based on an exhaustive analysis of the case of linear habit formation and linear Engel curves. Second, even when the long-run utility function exists, it is not an appropriate welfare criterion. The long-run utility function, when it exists, does not reflect "long-run preferences," but is merely an indicator of long-run behavior.

References

8. P. A. Samuelson, Social indifference curves, Quart. J. Econ. 70 (1956), 1–22.