Integrability, Generalized Separability, And a New Class of Demand Systems

Thibault Fally*
UC Berkeley ARE
April 2020

Abstract

This paper examines demand systems where the demand for a good depends on other prices only through a common price aggregator (a scalar function of all prices). This generalizes directly-separable preferences where the Lagrange multiplier provides such an aggregator. As indicated by Gorman (1972), symmetry of the Slutsky substitution terms implies that such demand can take only one of two simple forms, with either flexible price effects or flexible income effects. Furthermore, allowing for indirect utility as an additional aggregator generates a large variety of demand systems, some of them new, with greater flexibility in both price and income effects. Conversely, only weak conditions ensure that such demand systems can be rationalized, i.e. can be derived from the maximization of a well-behaved utility function. This paper provides examples and applications of these demand systems. In particular, they can be used in simple general-equilibrium models to generate a wide range of effects of market size and productivity on firm size, entry and prices.

Keywords: Consumer Demand, Separability, Price aggregator, Integrability, Rationalization, Non-homothetic preferences.

JEL Classification: D11, D40, L13

*University of California Berkeley, Agricultural and Resource Economics Department, NBER and CEPR. Many thanks to Philippe Aghion, Dave Donaldson, Federico Etro, Ben Faber, Ethan Ligon, Kiminori Matsuyama, Marti Mestieri, Andres Rodriguez-Clare, Ina Simonovska and Brian Wright for helpful comments and suggestions. Contact: fally@berkeley.edu
1 Introduction

The integrability problem, which consists in characterizing demand systems that can be rationalized and derived from utility functions, has long been a central issue in economic theory. Earliest contributions date from Antonelli (1886), with applications to various fields, including micro and macroeconomics, econometrics, industrial organization and international trade. Theorists have provided broad sufficient and necessary conditions for demand patterns to be integrable, notably Hurwicz and Uzawa (1971), who provide conditions based on the Slutsky substitution matrix, which must be symmetric and negative semi-definite for all prices and income levels.

While very general, the Hurwicz and Uzawa (1971) integrability conditions lack practicality. Perhaps a consequence is that applied theorists and practitioners have often focused on less general cases to ensure both tractability and rationality. In particular, one often focuses on directly-separable or indirectly-separable preferences. An attractive feature of these preferences is that demand depends only on a few variables, namely consumer income, a good’s own price, and a single aggregator (scalar) that is itself a function of the vector of prices and income. Such an aggregator can be, for instance, a price index (e.g. with constant elasticity of substitution preferences) or the marginal utility of income (with directly-separable preferences).\(^1\) These preferences, however, have properties that may be undesirable and too restrictive in terms of income and price effects. For instance, direct separability implies that income elasticities and price elasticities are proportional across goods (“Pigou’s law”), a testable prediction that has been empirically rejected, e.g., by Deaton (1974).

This paper characterizes demand systems that are more general but retain a key practical property of the widely-used demand systems mentioned above: the existence of a price aggregator that is common for all goods, a feature that is useful to demand estimation, welfare analysis, applied models of monopolistic and oligopolistic competition, and many other applications.

The paper aims to make three contributions. A first objective is to provide functional forms of demand (i.e. necessary conditions) to satisfy Slutsky symmetry when demand for a good depends only on its own price, income, utility and/or a common price aggregator. Then, a second objective is to provide sufficient conditions for such functional forms of demand to be rational, i.e. such that they can be derived from a well-defined quasi-concave utility function. A third contribution is to provide various examples of such demand systems, including some that have not been previously discussed in the literature, and illustrate how functional forms

\(^1\)In models with symmetric demand across product varieties with an upper bound in marginal utility for each variety, there exists a finite reservation price (or choke price) that can also be used as a common price aggregator (see e.g. Arkolakis et al., 2019).
of demand determine market size effects on firm size and prices in a simple general-equilibrium model.

Following Pollak (1972), we refer to a demand system as “generalized separable” if demand for each good $i$ satisfies:

$$q_i = \tilde{q}_i(p_i/w, \Lambda)$$  \hspace{1cm} (1)

where $p_i$ refers to its price, $w$ consumer income (total outlays), and where $\Lambda$ is a scalar (aggregator) that is a function of all prices and income. A key property of such demand is that all cross-price effects go through $\Lambda$, a practical property for modeling and for estimation, as the rank of the cross-price substitution matrix is then just one.

In fact, such demand system can only take some particular functional forms in order to be integrable. Providing the sketch of a proof that we complete here, Gorman (1972, 1995) indicates that such demand system can take either of two main forms if we impose the Slutsky substitution matrix to be symmetric:

$$q_i = \frac{D_i(F(\Lambda)p_i/w)}{H(\Lambda)}$$ \hspace{1cm} (2)

$$q_i = A_i(\Lambda)(p_i/w)^{-\sigma(\Lambda)}$$ \hspace{1cm} (3)

where $D_i$, $F$ and $H$ are positive real functions and where, in both cases, $\Lambda$ is a scalar variable that adjusts so that the budget constraint is satisfied, and can thus be defined as an implicit function of prices and income (under several additional assumptions on differentiability and invertibility). These demand systems have rarely appeared in the applied literature so far in spite of their usefulness.

In the first case (equation 2), sufficient conditions for integrability are expressed as conditions on elasticities of functions $H$, $F$ and $D_i$, and ensure that demand $q_i$ is decreasing in the aggregator $\Lambda$ for any good $i$. We will refer to this case as a “Gorman-Pollak” demand system. It corresponds to directly-additive utility when the quantity shifter $H(\Lambda)$ is constant and indirectly-additive utility when the price shifter $F(\Lambda)$ is constant. This also generalizes the results of Matsuyama and Ushchev (2017) on homothetic single-aggregator demand, which corresponds to the special case where $F(\Lambda) = 1/H(\Lambda) = \Lambda$. In this more general demand system, income and price elasticities both depend on the functional form chosen for $D_i$, which can be very flexible; demand and price shifters $H(\Lambda)$ and $F(\Lambda)$ also influence income effects and depend flexibly on the price aggregator. However, this formulation still imposes tight

---

2There are other cases that can be ruled out under additional restrictions on price sensitivity.

3There are a few recent exceptions, including Bertoletti and Etro (2018a,b) for the first case, Comin et al. (2015) and Matsuyama (2015) for the second case with homogeneous $\sigma(\Lambda) = \sigma$. 

2
constraints on price and income effects, as it implies an affine relationship between price and income elasticities of demand across goods for a given consumer.

In the second case (equation 3), with common price elasticities across goods, the aggregator $\Lambda$ actually coincides with indirect utility $V$ (up to a one-to-one mapping). In that case, integrability requires that the demand shifters $A_i(\Lambda)$ increase quickly enough in $\Lambda$. While quasi-concavity is easy to obtain in this case, conditions for rationalization need to ensure that indifference curves do not cross and that utility is monotonically increasing in the consumption of each good. Notice that the price elasticity $\sigma(\Lambda)$ does not have to remain constant or monotonic across indifference curves, i.e., indifference curves can become flatter or more convex as income goes up. We will refer to that case as “generalized non-homothetic CES”. This second case features Allen-Uzawa substitution elasticities that do not vary across goods but may vary with the demand aggregator $\Lambda$, and generalizes implicitly-additive utility functions previously defined by (Comin et al., 2015) who impose a constant elasticity of substitution $\sigma(\Lambda) = \sigma$. Relative to Gorman-Pollak demand, this case allows for more flexible income patterns, but rigid price effects.

Limitations of both cases with a single aggregator, either on price or income effects, call for considering demand systems that allow for two aggregators. The previous results can be extended by showing that any demand systems that depend on two aggregators, an aggregator $\Lambda$ and indirect utility $V$ (where $\Lambda$ is a function of normalized prices $p/w$ and is implicitly determined by the budget constraint) must take the following functional form:

$$q_i = \frac{D_i(F(\Lambda,V)p_i/w, V)}{H(\Lambda,V)}$$

(4)

where real-valued functions $D_i$, $F$ and $H$ now also depend on indirect utility $V$ as a second argument. Conversely, mild sufficient conditions on these functions ensure that such demand system is rational, and I provide ways to characterize direct and indirect utility functions as implicit functions. This form of demand encompasses the previous forms (2) and (3) based on a single aggregator, and can be used to generate a variety of new and more general demand systems. While the first aggregator $\Lambda$ must still affect demand only through common price and quantity shifters $H$ and $F$ as in the first case described above, demand $D_i$ for each good $i$ can be a very flexible function of indirect utility $V$. In particular, not only Engel curves can vary flexibly across different goods, but price effects and substitution patterns can also depend flexibly on income (through utility $V$). Such demand also encompasses many examples of systems used in the literature such as demand from directly or indirectly implicitly-separable preferences, various examples of homothetic demand systems, and symmetric versions of other commonly-used demand systems.
The single and dual-aggregator demand systems described here yield various applications. They are particularly useful in the case of monopolistic competition. In the limit where each firm has a negligible market share, it chooses its price by taking other prices and quantities as given. It is then practical to have a single industry-wide indicator $\Lambda$ that uniquely determines the locus of the demand curve for a good with respect to its own price.\footnote{Recent work by Bertoletti and Etro (2018a) formalizes this insight with asymmetric demand and covers the Gorman-Pollak demand system as an example. See also Anderson et al. (2018) on aggregative games where $\Lambda$ could be used as an “aggregate”.}

With the first type of demand system, the generalized Gorman-Pollak form, the price aggregator $\Lambda$ can be interpreted as an index of tightness of the budget constraint,\footnote{Aggregator $\Lambda$ is implicitly determined by the budget constraint and is thus similar to the Lagrange multiplier under directly-separable preferences.} or alternatively as an index of the toughness of competition in a model with firms. A change in the aggregator can lead to a vertical and an horizontal shift of each demand curve, with flexible implications for markups depending on the shape of these demand curves. This type of preferences can be used to rationalize many examples drawn from Mrázová and Neary (2013), e.g. bi-power and inverse bi-power demand functions, and Bulow-Pfleiderer demand (Weyl and Fabinger, 2013). For instance, with iso-elastic functions $H$ and $D_i$, they coincide with the self-dual addilog demand systems (Houthakker, 1965) and extend the constant relative income elasticities (CRIE) used for instance in Fieler (2011) and Caron et al. (2014). The functional form can also generate choke prices (as demand $D_i$ for a good $i$ goes to zero) which can be expressed as a simple function of income and the price aggregator, with a functional form that is again more flexible than commonly used in macroeconomics and international trade. In particular, this form can be used to generalize the results of Bertoletti and Etro (2017) and Bertoletti et al. (2018) in which the choke price is proportional to income (see Fally, 2019).

The second type of demand, generalized non-homothetic CES, is particularly relevant for situations where we want to allow for flexible income effects while retaining the most simple price effects. For instance, Comin et al. (2015) uses such demand with constant elasticity of substitution to model structural change and sector-specific Engel curves across agriculture, manufacturing and services. Atkin et al. (2018) use a similar demand structure to estimate welfare and price indices from shifts in Engel curves. With an elasticity of substitution that depends on utility, these preferences remain very tractable and empirically relevant. For instance, several studies (such as Handbury, 2013 and Faber and Fally, 2017) based on expenditure surveys and scanner data have shown that price elasticities vary significantly with income. Handbury (2013) and Faber and Fally (2017) model income effects in the elasticity of substitution by relying on a numeraire good. With generalized non-homothetic CES, we can instead generate such a relationship between income and the elasticity of substitution through utility, without
relying on a numeraire good.

The third and most general form with two aggregators can also be used for a variety of applications. A first motivation is to provide flexible demand systems for both price and income effects. Conversely, it is easy to manipulate and obtain desired properties that could be useful in specific settings. For instance, one can impose homotheticity while retaining very flexible price and substitution effects. Even in the homothetic case with unitary income elasticities, such demand system encompasses many examples that have been used in the literature, e.g. QMOR when $D_i$ is quadratic (Diewert, 1976; Feenstra, 2018), HDIA when $H$ is constant (Kimball, 1995; Matsuyama and Ushchev, 2017), HDIA when $F$ is constant, as well as HSA mentioned above when $\Lambda$ is the only aggregator (Matsuyama and Ushchev, 2017). Such demand system can also be used to rationalize the two-aggregator demand considered in Arkolakis et al. (2019), which is particularly appealing for its tractability and its applications to international trade models with heterogeneous firms. As shown by Thisse and Ushchev (2016), this demand system can be generated by aggregating over many rational consumers with random utility; here such demand is rationalized with a single representative consumer.

With a simple general-equilibrium model of homogeneous firms with monopolistic competition and free entry, a wide range of comparative statics can be qualitatively obtained with the Gorman-Pollak single-aggregator demand system. Even with a single aggregator, an increase in competition can shift demand curves both vertically and horizontally, and affect price elasticities by inducing a move along demand curves. A similar variety of outcomes can be obtained with implicitly-additive preferences. On the contrary, some types of separability (dubbed “semi-separability”) play a specific role in this framework by restricting the range of comparative statics. With directly-semi-separable preferences, firm output and prices do not depend on income (productivity) and only depend on population, while the number of firms is proportional to population. With indirectly-semi-separable preferences, firm output and prices do not depend on population and only depend on income. In the homothetic case, output, prices and entry only depend on total GDP. This analysis fits within the framework proposed by Parenti et al. (2017) based on very general symmetric demand systems, and also complements the results of Zhelobodko et al. (2012) based on directly-additively separable preferences, Bertoletti and Etro (2017) on indirectly-additively separable preferences, and Bertoletti and Etro (2018b) on Gorman-Pollak demand. This illustrates the importance of the choice of functional forms of demand in determining key outcomes in general-equilibrium models.

The paper relates to several others studying functional forms of utility and demand systems, with applications to demand estimation. In particular, Ligon (2016) focuses on cases where

---

6QMOR refers to: quadratic mean of order r expenditure function; HDIA: homothetic directly implicitly additive preferences; HIIA: homothetic indirectly implicitly additive; HSA: homothetic single aggregator.
the aggregator corresponds to the Lagrange multiplier \( \lambda \) associated with the budget constraint, and shows that a form of separability in \( \lambda \) implies specific functional forms as well as direct additive separability. Nocke and Schutz (2017) study the ("quasi-") integrability of quasi-linear demand systems, i.e. without income effects. Fabinger and Weyl (2016) examine functional forms of demand and production functions that lead to closed-form solutions in models imposing relationships between marginal and average effects. The discussion on the existence of aggregators also mirrors the restrictions associated with the rank of a demand system (Gorman, 1981; Lewbel, 1991; LaFrance and Pope, 2006; Lewbel and Pendakur, 2009). The rank of a demand system corresponds to the number of vectors and homothetic price aggregators needed to recover Engel curves. Here, the single aggregator \( \Lambda \) is generally not homogeneous of degree one in prices (and also depends on income) and the demand systems studied here do not have restrictions in terms of rank. Finally, Blackorby et al. (1978) study functional forms implied by various definitions of separability, and find that the same functional structure as with generalized non-homothetic CES is obtained when imposing stronger forms of separability that imply equality among Allen-Uzawa elasticities of substitution.

The remainder of the paper proceeds as follows. Section 2 examines the functional forms imposed by generalized separability with one and two aggregators. Section 3 provides sufficient conditions for each type of demand to ensure that it can be rationalized. Section 4 discusses various examples of these demand systems. Section 5 examines an application to monopolistic competition and studies market size effects in a simple general-equilibrium model.

2 Functional Forms under Generalized Separability

2.1 Single aggregator

Additively-separable utility allows us to obtain demand as a simple function of a good’s own price \( p_i \) and a single aggregator, the Lagrange multiplier. While practical, both direct and indirect separability put strong constraints on the structure of demand, such as a tight relationship between price elasticity and income elasticity, with for instance the adverse consequence that preferences with constant elasticity of substitution (CES) are the only directly-separable and indirectly-separable preferences that are homothetic.

In an attempt to generalize the concept of separability, Gorman (1972) and Pollak (1972)
define generalized separability as demand that would take the form:

\[ q_i = \tilde{q}_i(p_i/w, \Lambda) \]  

(5)

where demand for good \( i \) (quantity) is a real function of its own normalize price and the aggregator \( \Lambda \), i.e. a mapping \( \tilde{q}_i \) from \( \mathbb{R}_+ \times \mathbb{R}_+ \) to \( \mathbb{R}_+ \), and where \( w > 0 \) refers to total consumer expenditures and \( p_i > 0 \) refers to the price of good \( i \). \( \Lambda = \Lambda(p/w) \) a real function of the vector of normalized prices \( p/w = (p_1/w, ..., p_N/w) \in \mathbb{R}_+^N \), and \( N \in \mathbb{N} \) denote the number of goods. Without loss of generality, we assume that \( \Lambda \) is always positive.

We assume that the budget constraint holds for any vector of normalized prices \( p/w \), which implies that the aggregator \( \Lambda \) must satisfy:

\[ \sum p_i q_i = p_i \tilde{q}_i(p_i/w, \Lambda(p/w)) = w. \]

Under the regularity assumption [A1]-i) made below, the solution to this equation in \( \Lambda \) is unique and we can use the budget constraint to obtain the derivatives of \( \Lambda \) w.r.t. prices. Note that, generally, \( \Lambda \) is not a Lagrange multiplier, except for the case where demand can be derived from a directly-additive separable utility (Ligon 2016).

We say that the system of demand given by \( \tilde{q}_i \) and \( \Lambda \) is integrable if there exists a differentiable utility function \( U(q) \) such that marginal utility \( \frac{\partial U}{\partial q_i} \) evaluated at \( \tilde{q}_i \) (for a given vector of prices and income) is proportional to prices \( p_i \) across goods \( i \). We further assume that utility \( U \) is twice continuously differentiable, so that its cross-derivatives are symmetric.

For the sake of simplicity and exposition, we focus on demand that can be inverted and assume that for each vector \( q \in \mathbb{R}_+^N \), there exists a vector \( p/w \in \mathbb{R}_+^N \) such that \( q_i = \tilde{q}_i(p_i/w, \Lambda(p/w)) \).

In an unpublished note by Gorman (printed in Gorman, 1995) mentioned by Pollak (1972), Gorman indicates that a demand system defined as above needs to take specific forms in order to satisfy Slutsky’s symmetry condition. With a few additional restrictions, this result can be formulated as follows:

---

8A distinction is often made between integrability and rationalization, whereby the latter further requires \( U \) to be quasi-concave. In other words, integrability imposes the Slutsky substitution matrix to be symmetric while rationalization also requires that is is semi-definite negative.

9This assumption is made for convenience as the proof mostly focuses on the inverse demand. Most of the arguments are local and would apply to subsets of prices and quantities where we have invertibility.

10Gorman’s sketch of proof had many shortcuts, as he himself noted: “Throughout this paper I have talked as if my claims were definitely proven. Of course this is not so: my arguments are far from rigorous” (Gorman, 1995). Here I impose somewhat stronger assumptions on the form of demand and price effects in order to avoid a few inelegant cases. In particular, the assumption that expenditure shares are not just a function of \( \Lambda \) allows me to avoid what Gorman calls “the abnormal case”.

7
Regularity assumptions [A1] on functions $\tilde{q}_i$:

i) $\tilde{q}_i(p_i/w, \Lambda)$ is positive and twice continuously differentiable, with strictly negative derivatives in both arguments;

ii) Holding $\Lambda$ constant, $p_i\tilde{q}_i(p_i/w, \Lambda)$ has a non-zero derivative in $p_i$

iii) There are at least four goods ($N \geq 4$);

iv) Invertibility: for each $q_i \in \mathbb{R}_+^N$ there exists $p_i/w \in \mathbb{R}_+^N$ such that $q_i = \tilde{q}_i(p_i/w, \Lambda(p_i/w))$ for all $i$.

Proposition 1 If demand is integrable and satisfies conditions [A1], it can be written as either:

**case 1:**

$$\tilde{q}_i(p_i/w, \Lambda) = \frac{D_i(F(\Lambda)p_i/w)}{H(\Lambda)}$$

for all goods $i$ and all $p_i, w, \Lambda$

**case 2:**

$$\tilde{q}_i(p_i/w, \Lambda) = A_i(\Lambda)(p_i/w)^{-\sigma(\Lambda)}$$

for all goods $i$ and all $p_i, w, \Lambda$

**case 2':**

$$\tilde{q}_i(p_i/w, \Lambda) = a_iA(\Lambda)(p_i/w)^{-\sigma_0}$$

for all but one good $i$

or a combination of cases 2 and 2' (depending on $\Lambda$), where $a_i$, $\sigma_0$ and $\rho_0$ are positive constant terms, and $D_i$, $F$, $H$, $A$ and $A_i$ are differentiable real functions with a single argument.

To prove Proposition 1, it is actually easier to work with the inverse demand, i.e. expressing normalized prices as a function of quantities, as such object is more directly related to marginal utility.\(^{11}\) A key integrability condition comes from the symmetry of the Hessian of the utility function. As inverse demand is proportional to marginal utility, its derivatives also need to feature some symmetry, a condition equivalent to Slutsky symmetry for Marshallian demand.\(^{12}\) When cross-price effects are captured by a single aggregator, these symmetry conditions impose conditions on price elasticities that can then be integrated to provide the functional forms in Proposition 1. As part of the proof of Proposition 1, we can show that inverse demand takes a very similar functional form in both cases. In the first case, we can express inverse demand as:

$$\frac{p_i}{w} = \frac{D_i^{-1}(H(\Lambda)q_i)}{F(\Lambda)}$$

where $\Lambda$ is now seen as a function of the vector of consumption and can be implicitly defined as a solution to the budget constraint using inverse demand: $\sum_i q_i D_i^{-1}(H(\Lambda)q_i)/F(\Lambda) = 1$.\(^{13}\)

---

\(^{11}\)A working paper version provides a proof based on direct demand, examining price effects and Slutsky symmetry, which more closely follows the steps proposed by Gorman (1972) in his unpublished notes.

\(^{12}\)Earlier literature on rationalization (e.g. Samuelson, 1950) put more attention onto properties of inverse demand than relatively more recent works that have placed a greater emphasis on the Slutsky substitution matrix (following Hurwicz and Uzawa, 1971).

\(^{13}\)In the Marshallian demand formulation, $\Lambda$ is a function of prices and income. In the inverse demand formulation, we redefine $\Lambda$ as a function of quantities, where $\Lambda$ can again be implicitly characterized by the budget constraint. As an abuse of notation, we use the same notation in both approaches.
This inverse demand formulation highlights the symmetric role of $H$ and $F$.

Since the third case 2' is relatively less interesting and elegant (CES for all but one good), the remainder of the paper focuses on cases 1 and 2, setting aside case 2'. Note that there may be alternative functional forms under generalized separability if we allow for price-insensitive expenditures shares, which Gorman calls "abnormal" goods. Assumption iii) allows us to exclude such cases. Also note that functional forms are unique up to a constant term and a monotonic transformation of $\Lambda$, both in cases 1 and 2. Moreover, as we will see later in Section 3.2 under additional restrictions, the aggregator $\Lambda$ coincides in case 2 with indirect utility $V$, up to a monotonic transformation.

Before turning to demand with two aggregators and sufficient conditions for rationalization, it is useful at this point to summarize some key properties implied by these two types of functional forms of demand, especially in terms of price and income effects.

**Price and income elasticities in case 1.** Let us denote by $\varepsilon_{Di} = \frac{\partial \log D_i}{\partial \log p_i}$, $\varepsilon_H = \frac{\partial \log H}{\partial \log \Lambda}$ and $\varepsilon_F = \frac{\partial \log F}{\partial \log \Lambda}$ the elasticity of $D_i$, $H$ and $F$ in their argument. In case 1, the price elasticity of Marshallian demand is:

$$\frac{\partial \log q_i}{\partial \log p_j} = \varepsilon_{Di} \cdot 1_{(i=j)} - \frac{W_j(1 + \varepsilon_{Dj})(\varepsilon_H - \varepsilon_F \bar{\varepsilon}_D)}{\varepsilon_H - \varepsilon_F \bar{\varepsilon}_D}$$

(7)

where $W_j$ is the expenditure share of good $j$, $1_{(i=j)}$ is a dummy equal to one when $i = j$, and $\bar{\varepsilon}_D = \sum_i W_i \varepsilon_{Di}$. When that good has a negligible market share, the own price elasticity is determined by the shape of function $D_i$: $\frac{\partial \log q_i}{\partial \log p_i} \approx \varepsilon_{Di}$. Since we impose few constraints on $\varepsilon_{Di}$, the shape of each demand curve and the patterns of price elasticities can be very flexible.

In turn, the income elasticity of demand is:

$$\frac{\partial \log q_i}{\partial \log w} = 1 + \frac{(\varepsilon_H + \varepsilon_F)(\bar{\varepsilon}_D - \varepsilon_{Di})}{\varepsilon_H - \varepsilon_F \bar{\varepsilon}_D}.$$  

(8)

Using this expression, one can see that homotheticity implies that either $\varepsilon_H = -\varepsilon_F$ or $\varepsilon_{Di} = \bar{\varepsilon}_D$ for all $i$ (see Section 4.2).

As pointed out by Pigou (1910) and Deaton (1974), own-price elasticities and income elasticities are tightly linked (across goods) when demand is derived from a directly-additive utility, which corresponds to the case where $\varepsilon_H = 0$. With direct additive separability and negligible expenditure shares $W_i$, we obtain that the two are proportional across goods: $\frac{\partial \log q_i}{\partial \log w} = \frac{\varepsilon_{Di}}{\varepsilon_D}$. When $\varepsilon_H \neq 0$, the relationship between income elasticity and price elasticity is muted and is now affine. The relationship can also be flipped if $\varepsilon_F + \varepsilon_H > 0$, with price-elastic goods being relatively less income elastic.
Price and income elasticities in case 2. In the second case, price effects are simpler: the own-price elasticity is constant for a given level of the aggregator $\Lambda$, and it is constant across all goods in the limit case where each good has a negligible expenditure share. As we will see in Proposition 3, in case 2 we can generally interpret the aggregator $\Lambda$ as indirect utility, and thus $\sigma(\Lambda)$ also corresponds to the price elasticity of Hicksian demand.

This demand system is most interesting and useful for its very flexible income effects. Comparing goods, first we can see that changes in $G_i(\Lambda)$ in $\Lambda$ need not be related to changes in $\sigma(\Lambda)$, thus breaking away from the link between price and income elasticities discussed for the first case above. Starting with the special case where $\sigma(\Lambda) = \sigma$ is constant, the effect of income on $\Lambda$ is such that:

$$\frac{\partial \log \Lambda}{\partial \log w} = 1 \bar{\epsilon}_A$$

(9)

where $\bar{\epsilon}_A$ is an average of elasticities $\epsilon_{Ai} = \frac{\Lambda A_i'(\Lambda)}{A_i(\Lambda)}$ weighted by expenditures shares. We obtain the income elasticity of demand:

$$\frac{\partial \log q_i}{\partial \log w} = \frac{\epsilon_{Ai}}{\bar{\epsilon}_A}. \tag{10}$$

Hence income effects are captured by how the shifters $A_i$ vary with $\Lambda$. In particular, good $i$ is income-elastic if and only if $\epsilon_{Ai}/\bar{\epsilon}_A > 1$. In the more general case where $\sigma(\Lambda)$ is not constant, function $A_i$ plays a similar role and dictates income effects, while $\sigma(\Lambda)$ determines how the price elasticity varies with income.

2.2 Generalization with two aggregators

As we have seen above, the single-aggregator case either imposes a tight constraint on price elasticities (case 2) or an affine relationship between price and income elasticities (case 1), and excludes several demand systems examined in the literature which more generally require two aggregators. The next objective is to examine how to combine case 1 and case 2 above, by studying demand systems that depend on two aggregators, one of them coinciding with utility as in case 2. For various applications, it is useful to model both flexible price effects and flexible income effects—that are not as tightly linked as with a single aggregator—and provide a general formulation that includes most demand systems used in practice.

Hence, we now suppose that demand takes the form:

$$q_i = \tilde{q}_i(p_i/w, \Lambda, V) \tag{11}$$

where $\tilde{q}_i$ is now a mapping from $\mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}$ to $\mathbb{R}_+$, where $V = V(p/w)$ refers to the indirect
utility function evaluated at \( p/w \) and where \( \Lambda \) satisfies the budget constraint:

\[
\sum_i \frac{p_i q_i}{w} = \sum_i \frac{p_i \tilde{q}_i(p_i/w, \Lambda, V)}{w} = 1.
\]

Under conditions \([A2]\) imposed for Proposition 2, inverse demand \( \tilde{q}_i^{-1} \) is well defined and can be expressed as a function of quantity \( q_i \), the direct utility function \( U \) and the aggregator \( \Lambda \), which we can alternatively express as a function of quantities \( q \) such that the budget constraint holds (see Appendix for details). This gives:

\[
p_i/w = \tilde{q}_i^{-1}(q_i, \Lambda, U) \tag{12}
\]

where \( \tilde{q}_i^{-1} \) denotes the inverse demand w.r.t. normalized prices, and where \( \Lambda = \Lambda(q) \) is now implicitly defined as a function of quantities. We use the fact the derivatives of the indirect utility function and the direct utility function are proportional to demand \( q_i \) across goods \( i \).

We can generalize Proposition 1 under a set of similar regularity restrictions on differentiability, minimum number of goods and price effects:

**Regularity assumptions** \([A2]\) on functions \( \tilde{q}_i \):

i) \( \tilde{q}_i(p_i/w, \Lambda, V) \) is positive and twice continuously differentiable, with a strictly negative derivative in \( p_i \) and \( \Lambda \);

ii) Holding \( \Lambda \) and \( V \) constant, \( p_i \tilde{q}_i(p_i/w, \Lambda, V) \) is not constant over the range of prices \( p_i \);

iii) There are at least four goods and, for any vector of normalized prices \( p/w \) (except for a set of prices of measure zero), the price elasticity takes at least three values across goods.

iv) Invertibility: for each \( q \in \mathbb{R}^N_+ \), \( \exists p/w \in \mathbb{R}^N_+ \) s.t. \( q_i = \tilde{q}_i(p_i/w, \Lambda(p/w), V(p/w)) \) for all \( i \).

**Proposition 2** If demand \( \tilde{q}_i \) is integrable, depends on two aggregators as in equation (11) and satisfies regularity conditions \([A2]\), it can be written as:

\[
\tilde{q}_i(p_i/w, \Lambda, V) = \frac{1}{H(\Lambda, V)} D_i \left( \frac{p_i F(\Lambda, V)}{w}, V \right) \tag{13}
\]

where \( D_i, F \) and \( H \) are mappings from \( \mathbb{R}_+ \times \mathbb{R} \) to \( \mathbb{R} \), with indirect utility \( V \) as second argument.

This functional form is again imposed by symmetry conditions. The proof of Proposition 2 follows similar steps as for Proposition 1. Surprisingly, these symmetry conditions do not impose strong constraints on functional forms in terms of how indirect utility, used as a second
aggregator, influences demand patterns. This leads to much more flexibility: indirect utility can influence partial demand functions $D_i$ as a second argument, in a way that can be specific to each good, and can also enter the price and quantity shifters, functions $F$ and $H$, in addition to the first aggregator $\Lambda$. Such functional form provides a generalization of both cases 1 and 2 that we encounter when we impose a single-aggregator $\Lambda$.

In terms of inverse demand, the functional form is quite similar as we obtain:

$$\tilde{q}_i^{-1}(q_i, \Lambda, U) = \frac{1}{F(\Lambda, U)} D_i^{-1}(q_i H(\Lambda, U), U)$$

where $D_i^{-1}$ denotes the inverse of $D_i$ with respect to its first argument, and where $U$ refers to direct utility.

In order to avoid a taxonomy of cases, condition [A2]-iii) above imposes enough heterogeneity in price elasticities across goods. This excludes special cases that would resemble cases 2 and 2’ in Proposition 1. This restriction does not lead to an important loss of generality given that a key motivation for Proposition 2 is to examine more flexible demand systems in both income and price effects.

**Price and income effects.** In the single-aggregator case, we have seen that the own-price elasticity is given by the elasticity of $D_i$, and thus the shape of $D_i$ influences how price elasticities (and thus markups in models of imperfect competition) vary along the demand curve depending on the level of demand for a particular good $i$. With utility $V$ as an additional aggregator, the shape of demand curves can itself vary with utility, the second argument of $D_i$, and therefore vary across income levels. This generalizes the effect of utility on $\sigma$ in case 2 described in the previous section (CES case), now with a demand curve that does not have to be iso-elastic.

In case 1 with a single aggregator, all changes in other prices (e.g. a change in competition) are captured by $\Lambda$, which then influences how demand curves shift vertically and horizontally, depending on the demand shifters $F$ and $H$. With two aggregators, an change in overall welfare can more directly affect these demand shifters (through $V$), and other effects of prices captured by $\Lambda$ can now also vary across utility levels, and thus across income levels. For instance, this could be used to model more subtle interactions between per capita income and the degree of competition in a market.

How these price and competition effects vary with utility can remain quite flexible. At this point, we only impose restrictions on Slutsky symmetry (integrability), but we will see in the next section that fairly mild additional restrictions are sufficient to ensure that such systems are rational, so these considerations on price and income effects will remain valid.
3 Rationalization

Let us now examine the reciprocals of Proposition 1 and 2. Under which conditions these demand systems can be rationalized, i.e. can be derived from maximizing a well-behaved quasi-concave and monotone utility function? These functional forms, imposed by the symmetry of the Slutsky matrix, do not necessarily lead to quasi-concavity or monotonicity of a utility function (see Appendix for counter-examples in the single-aggregator case). However, in each case, it turns out that only weak and simple additional conditions are sufficient to guarantee that the demand systems described in Proposition 1 and 2 are rational.

3.1 Rationalization of Gorman-Pollak Demand

Suppose that demand is given by:

\[ q_i = \frac{D_i(F(\Lambda)p_i/w)}{H(\Lambda)} \]  

(14)

where \( D_i, F \) and \( H \) are mappings from \( \mathbb{R}_+ \) to \( \mathbb{R}_+ \), and where \( \Lambda \) is implicitly determined by the budget constraint \( \sum_i p_i D_i(F(\Lambda)p_i/w)/H(\Lambda) = w \), which can be rewritten:

\[ H(\Lambda) = \sum_i (p_i/w) D_i(F(\Lambda)p_i/w). \]  

(15)

As before, we denote by \( \varepsilon_{Di} = \frac{\partial \log D_i}{\partial \log p_i} \) the elasticity of \( D_i \) in its argument, and \( \varepsilon_F = \frac{\partial \log F}{\partial \log \Lambda} \) and \( \varepsilon_H = \frac{\partial \log H}{\partial \log \Lambda} \) the elasticity of \( F \) and \( H \) in \( \Lambda \). To ensure that (15) has a unique solution in \( \Lambda \) and that this demand system is well-defined and rational, we impose the following regularity restrictions on \( D_i, F \) and \( H \):

Regularity assumptions [A3] on functions \( D_i, F \) and \( H \):

i) \( D_i \) is continuously differentiable, \( \varepsilon_{Di} < 0 \);

ii) \( H \) and \( F \) are continuously differentiable and \( \varepsilon_F \varepsilon_{Di} < \varepsilon_H \) for all \( i, \Lambda \) and \( p_i/w \)

iii) For any good \( i \) and \( y_i > 0 \), there exists \( \Lambda \in \mathbb{R}_+ \) such that: \( y_i D_i(y_iF(\Lambda))/H(\Lambda) = 1/N \)

Note that instead of condition [A3]-ii) we could assume that \( \varepsilon_F \varepsilon_{Di} - \varepsilon_H \) has the same sign for all \( i, \Lambda \) and \( p_i/w \). Assuming that this difference is negative is without loss of generality as we can always make the change in variable \( \Lambda' = 1/\Lambda \) to switch the sign of this inequality for all goods and prices. Assumptions i) and ii) imply that the solution in \( \Lambda \) to equation (15) is always unique, but they are also needed to show that utility is quasi-concave and that the Slutsky substitution matrix is negative semi-definite. Condition iii) ensures that equation (15)
has a solution in Λ: in other words, the aggregator Λ can always adjust in order to satisfy the budget constraint. Condition iii) is automatically satisfied, for instance, if we assume that the image of the mapping $\Lambda \mapsto \frac{D_i(F(\Lambda)p_i/w)}{H(\Lambda)}$ is $(0, +\infty)$, conditional on $p_i/w$.\(^{14}\)

As with Proposition 1, it is useful to consider the inverse demand, which shares a similar functional form. We can define Λ as an implicit function of $q$ using inverse demand and the budget constraint:

$$\sum_i q_i D^{-1}_i(H(\Lambda)q_i) = F(\Lambda) \quad (16)$$

which, under conditions ii) and iii) has a unique solution in Λ for any $q$ (see Appendix).

Under these conditions, we obtain:

**Proposition 3** If $H$ and $D_i$ satisfy the regularity conditions [A3], the demand described in equations (14) and (15) can be rationalized and obtained from a continuous quasi-concave utility:

$$U(q) = \sum_i \int_{q' = q_{i0}}^{x = H(\Lambda(q))q_i} D^{-1}_i(q')dq' - \int_{l = \Lambda_0}^{\Lambda(q)} H'(l)F(l)dl \quad (17)$$

where $\Lambda(q)$ satisfies (16) for each $q$, and $\Lambda_0, q_{i0} \geq 0$ are constant terms.

Proposition 3 rationalizes such demand in a constructive way, by directly providing a utility function (see Appendix for details).\(^{15}\) Note that this utility function is unique, up to a monotonic transformation. The least obvious part of the proof is to show that such utility is quasi-concave, accounting for how the aggregator Λ responds to changes in $q$. An alternative is to build on the proof provided by Matsuyama and Ushchev (2017) for the homothetic case, as their approach can be also be extended to the non-homothetic case. In brief, this alternative approach checks that the Slutsky substitution matrix is symmetric and semi-definite negative, so that we can apply Hurwicz and Uzawa (1971) theorem.

Note that equation (16) can be seen as a first-order condition such that the expression above

---

\(^{14}\)Alternatively, in condition iii), one could replace the term $1/N$ on the left hand side by a term that varies across goods $i$ as long as this term sums up to unity across goods.

\(^{15}\)This utility representation was pointed out by Gorman (1987) with a more restrictive formulation and no formal proof that such utility function is well defined and quasi-concave. Gorman formulated this as a maximization: $U = \max_\Lambda \{\sum_i u_i(\Lambda q_i) - \Phi(\Lambda)\}$ but this approach is equivalent to assuming $H'(\Lambda) > 0$ in the formulation provided here, and omits very useful cases (such a continuum of cases providing a bridge between directly-additive and homothetic-single-aggregator preferences) where the second order condition of this maximization is not satisfied yet the utility function remains quasi-concave with Λ implicitly defined by equation (16).
for $U$ has a zero derivative in $\Lambda$. As such, marginal utility takes a simple form:

$$\frac{\partial U}{\partial q_i} = HD_i^{-1}(Hq_i).$$ \hfill (18)

Thanks to Proposition 1, we already know that it is symmetric but this does not ensure semi-definite negativity. As one could expect, the conditions ensuring the semi-definite negativity of the Slutsky matrix are the same as those providing the quasi-concavity of the utility function above.

Next, an important concern is whether the set of conditions [A3] can be relaxed, but I argue here that all are needed. First, the demand system would clearly not be well defined if it does not have a solution in equation (15), so condition iii) is unavoidable. It is possible to impose simpler conditions to ensure existence, but such conditions would be less general or practical. Second, restriction ii) is the simplest and more direct way to ensure that the equation defining the price aggregator has a unique solution. It is required for good $i$ for a given level of prices when a good $i$ has a sufficiently large expenditure share. In the appendix, I provide an example with two goods where restrictions i) and iii) are met but the Slutsky matrix is no longer negative semi-definite when $\varepsilon_F \varepsilon_{Di} - \varepsilon_H$ does not have the same sign for the two goods. Finally, restriction i) ensures that we have a negative effect of prices on demand when the expenditure share of a good is small (a positive price effect would not be rational for small expenditure shares). Inverting $D_i$ is also needed in equations (16) and (17) to retrieve utility.

Drawing from Pollak (1972), indirect utility can be expressed as:

$$V(p, w) = -\sum_i \int_{y_i0}^{(p_i/w)F(\Lambda)} D_i(y)dy + \int_{\Lambda_0}^{\Lambda} F'(l)H(l)dl + g_0$$ \hfill (19)

where $y_{i0}$, $g_0$ and $\Lambda_0$ are constant terms (see details in Appendix). $\Lambda = \Lambda(p/w)$ can either be implicitly defined by the budget constraint as above, or by taking the derivative of expression (19) w.r.t. $\Lambda$. This expression can also be useful to compute equivalent and compensating variations, implicitly defined such that $V(p', w - CV) = V(p, w)$ and $V(p, w + EV) = V(p', w')$. Taking the derivative w.r.t. income, one can interpret the product of the two shifters as the marginal utility of income (in log):

$$\frac{\partial V(p, w)}{\partial \log w} = F(\Lambda)H(\Lambda).$$

In terms of price and income effects, already discussed in Section 2.1 (expressions 7 and 8), assumptions [A3] do not impose stark additional restrictions. Given that we assume $\varepsilon_H > \varepsilon_F \varepsilon_{Di}$,
note however that the own-price elasticity is always negative, which rules out Giffen goods. Given that restriction, we can also see that the cross-price elasticity \((i \neq j)\) is positive if and only if \(\varepsilon_{Dj} < -1\).

Such demand is slightly more general than the one used in Pollak (1972) and more recently in Bertoletti and Etro (2018a) as it does not require either \(F(\Lambda)\) and \(H(\Lambda)\) to be monotonic in \(\Lambda\). If \(F'(\Lambda) > 0\), an increase in \(\Lambda\) (tightness of the budget constraint) leads to a downward shift in the partial demand curve \(D_i\). When \(F'(\Lambda) < 0\), we would instead have an upward shift in \(D_i\), which needs to be compensated by a large enough decrease in the demand shifter \(H(\Lambda)\). If \(F(\Lambda)\) is strictly monotonic (which is satisfied in practice for most applications, see e.g. Fally, 2019), then without loss of generality we can assume \(F(\Lambda) = \Lambda^\beta\) with \(\beta \in \{-1, 0, 1\}\), and thus:

\[
q_i = D_i(\Lambda^\beta p_i/w)/H(\Lambda). \quad (20)
\]

### 3.2 Rationalization of Generalized Non-Homothetic CES

Now, consider the second case of Proposition 1. Let us assume that expenditure shares are given by:

\[
p_i q_i / w = (G_i(\Lambda)p_i/w)^{1-\sigma(\Lambda)} \quad (21)
\]

where \(\sigma\) and each \(G_i\) is a continuous mapping from \(\mathbb{R}_+\) to \(\mathbb{R}_+\), where \(\Lambda(p/w)\) is itself a function of the vector of normalized prices \(p/w\). We assume that the budget constraint is satisfied, i.e.:

\[
\sum_i (G_i(\Lambda)p_i/w)^{1-\sigma(\Lambda)} = 1. \quad (22)
\]

To ensure integrability, we impose the following sufficient regularity restriction \([A4]\):

**Regularity assumptions \([A4]\)** For each \(\Lambda\), we have \(\sigma(\Lambda) \neq 1\) and *either* one of the following two conditions:

i) \(\sigma(\Lambda)\) is weakly increasing in \(\Lambda\) and \(G_i(\Lambda)\) is strictly increasing in \(\Lambda\)

ii) \(\sigma(\Lambda)\) is decreasing in \(\Lambda\) and, for each \(\Lambda_0\), there exists \(\alpha_i > 0\) such that \(\sum_i \alpha_i = 1\) and such that \(G_i(\Lambda)^{\sigma(\Lambda)-1}\) is strictly increasing in \(\Lambda\) in a neighborhood of \(\Lambda_0\)

Continuity is sufficient for the main statement. However, when both \(\sigma(\Lambda)\) and \(G_i(\Lambda)\) are all differentiable, condition ii) can be rewritten after solving for the minimum \(\alpha_i\) that would satisfy
this monotonicity condition. Condition ii) is formally equivalent to imposing:

$$\sum_i \exp \left( \frac{\sigma(\Lambda) - 1}{\sigma(\Lambda)} G_i'(\Lambda) \right) < 1 \quad (23)$$

(see Appendix for the proof of equivalence).

Under these conditions, we obtain the following proposition for the generalized case of non-homothetic CES:

**Proposition 4** Suppose that demand can be written as in equation (21) where $G_i$ and $\sigma$ are continuous and where $\Lambda$ is implicitly defined by (22). This demand system is integrable if conditions [A4] are satisfied. Under [A4], demand can be derived from a utility function that is implicitly defined by:

$$\sum_i \left( \frac{q_i}{G_i(U)} \right)^{\sigma(U)/\sigma(\Lambda) - 1} = 1 \quad (24)$$

which has a unique solution in $U$, with $\Lambda = U$ for the demand $q_i$ described above.

The constant elasticity case $\sigma(\Lambda) = \sigma$ corresponds to implicitly additive utility as in Comin et al. (2015). This is not equivalent to the standard CES since, even in that case, non-trivial income effects through the demand shifter $G_i(\Lambda)$ allow for very flexible Engel curves. The main contribution of this proposition is to generalize to variable elasticity of substitution.

The proof of Proposition 4 mainly consists in showing that $\Lambda$ is well-defined, i.e. that the budget constraint has a unique solution in $\Lambda$, and that utility is also uniquely defined by equation (24). As the more general case allows for varying curvature of indifference curves, one needs to ensure in particular that these indifference curves do not cross.

The proof proceeds as follows. First we show in a lemma that $\left[ \sum_i \alpha_i x_i^{\rho} \right]^{1/\rho}$ is monotonically increasing in $\rho$ if $\sum_i \alpha_i = 1$ (a consequence of Jensen’s inequality). This allows us to obtain comparative statics in the exponent in equations (24) and (22). We can then show that the solutions to these equations are unique, for a given set of income and prices, or quantities. Once we have uniqueness, it is easy to verify the quasi-concavity of the utility function (as in Comin et al., 2015). The last step is to check that this utility maximum problem does yield the demand system described above.

Again, as for Proposition 3, a potential concern is whether restrictions [A2] are necessary. When neither condition i) or ii) is satisfied, neither the demand system described above nor

---

16In general, note that condition ii) need not hold for *any* set of $\alpha_i$’s, it is sufficient that it holds for a single set of $\alpha_i$’s. In particular, using $\alpha_i = 1/N$ (where $N$ denotes the number of goods), a sufficient condition is that $G_i(\Lambda) N^{\frac{1}{\sigma(\Lambda)}}$ strictly increases in $\Lambda$. 

17
the utility in Proposition 3 is well defined. Counter-examples in the appendix further illustrate
the role of each condition, showing that equations (22) and (24) admit multiple solutions in \( \Lambda \)
and \( U \) if conditions i) and ii) are not satisfied. Incidentally, this shows that monotonicity in
demand shifters \( G_i(\Lambda) \) is not sufficient.\(^{17}\)

One should also point out why we need different conditions depending on whether \( \sigma(\Lambda) \)
decrees or increases with \( \Lambda \). In the first case, where \( \sigma(\Lambda) \) increases with \( \Lambda \), indifference
curves become flatter as we move away from the origin (with increases in income and \( \Lambda \)). In
that case, indifference curves are most likely to cross around the intercepts (when only one
good is consumed). Monotonicity in \( G_i(\Lambda) \) is then sufficient to ensure that indifference curves
do not cross. In the second case, where the elasticity of substitution \( \sigma(\Lambda) \) decreases with \( \Lambda \), the
indifference curves are more curved as we move away from the origin. In this case, indifference
curves are most likely to be close to each other and intersect around their midpoint.

3.3 Rationalization with two aggregators

Suppose that demand takes the form:

\[
q_i(p_i/w, \Lambda, V) = \frac{1}{H(\Lambda, V)}D_i\left(\frac{p_i F(\Lambda, V)}{w}, V\right)
\]  

(25)

where \( D_i \), \( F \) and \( H \) all are positive continuously-differentiable mappings from \( \mathbb{R}_+ \times \mathbb{R} \) to \( \mathbb{R}_+ \), with
aggregator \( V \) as second argument (which must coincide with indirect utility if such demand is
rational). Denote by \( \varepsilon_{D_i} \) the elasticity of \( D_i \) with respect to price \( p_i \) (holding \( \Lambda \) and
\( V \) constant) and by \( \varepsilon_H \) and \( \varepsilon_F \) the elasticities of \( H \) and \( F \) in terms of \( \Lambda \) (holding \( V \) constant).

We then impose the following sufficient regularity restrictions:

**Regularity assumptions** [A5] on functions \( D_i, F \) and \( H \):

i) \( D_i \) is continuously differentiable, with \( \varepsilon_{D_i} < 0 \);

ii) \( H \) and \( F \) are continuously differentiable, with \( \varepsilon_F \varepsilon_{D_i} < \varepsilon_H \) for all \( i, \Lambda, V \) and \( p_i/w \);

iii) For any good \( i, y_i > 0, V \in \mathbb{R}, \exists \Lambda \in \mathbb{R}_+ \) such that: \( y_i D_i(y_i F(\Lambda, V), V)/H(\Lambda, V) = 1/N \).

These conditions ensure that for each \( V \) and \( p/W \), there is a unique \( \Lambda \) such that the budget
constraint is satisfied, i.e. such that \( \sum_i(p_i/w)q_i(p_i/w, \Lambda, V) = 1 \) with demand defined in
equation (25) above. A similar result is obtained for the inverse demand. For any given vector
of quantities \( q \) and utility \( U \), the following budget condition for inverse demand:

\[
\sum_i q_i D_i^{-1}(q_i H(\Lambda, U), U)/F(\Lambda, U) = 1
\]  

(26)

\(^{17}\)We can also have \( \sigma(\Lambda) = 1 \) for a discrete number of values of \( \Lambda \).
has a unique solution in $\Lambda$.

These conditions are very similar to those used in the single-aggregator case for Gorman-Pollak demand in Proposition 3.\footnote{Again, as in Proposition 3, in condition iii) one could replace the term $1/N$ by a series of good-specific terms that sum up to unity across goods.} Under these conditions, we obtain the following proposition characterizing utility for more general demand systems with two aggregators including utility:

**Proposition 5** Suppose that demand can be written as in equation (25) satisfying regularity assumptions [A5] above, where $V$ with indirect utility and $\Lambda$ is an aggregator such that the budget constraint (26) holds. Then:

i) Utility $U$ must satisfy:

$$
\sum_i \int_{q=q_i0}^{q_iH(\Lambda,U)} D_i^{-1}(q,U) dq - G(\Lambda,U) = 0
$$

for some constant terms $q_i0 \geq 0$ and a continuously differentiable real-valued function $G(\Lambda,U)$ such that $\frac{\partial G}{\partial \Lambda}(\Lambda,U) = \frac{\partial H}{\partial \Lambda}(\Lambda,U)F(\Lambda,U)$.

ii) Conversely, if the left-hand-side of equation (27) is decreasing in $U$ (with a strictly negative partial derivative in $U$), equations (27) and (26) uniquely characterize a well-behaved utility (monotonic, continuous and quasi-concave) that yields demand as in equation (25).

Taken together, under conditions [A2] and [A5], Propositions 2 and 5 provide a characterization of rational demand functions with two aggregators $\Lambda$ and $V$ capturing cross-price effects, and a characterization of their associated utility functions.

The proof of Proposition 5 (see Appendix) combines elements of Proposition 3 and 4. First, the implicit solution for utility $U$ must be monotonically increasing in $q_i$ for each good $i$. Here, this property is obtained by assuming that the left-hand side of equation (27) is decreasing in $U$ (conditional on $q$ and $\Lambda$), given that the left-hand side has a strictly positive derivative in each $q_i$ and has a zero derivative in $\Lambda$. Proposition 5 does not provide a precise criteria, such as in Proposition 4, to determine when the left-hand side of equation (27) is decreasing in $U$, but in practical cases this condition is easy to check (as in the examples provided in Section 4).

Next, the proof that utility $U$ is quasi-concave in $q$ is similar to the one in Proposition 3 for the single-aggregator case. Considering the left-hand-side of equation (27) as a function of $q$ and $U$, it suffices to show that it is quasi-concave in $q$ (holding $U$ constant) in order to obtain that the implicit function for $U$ is quasi-concave in $q$. Holding $U$ constant, we can see that the
left-hand-side of equation (27) has the same structure w.r.t. \( q \) and \( \Lambda \) as the right-hand side of equation (17) for utility in the single-aggregator case in Proposition 3.

One must also ensure that \( \Lambda \) is well defined (implicitly defined such that the budget constraint holds). Condition [A5]-iii) leads to the existence of \( \Lambda \) while condition ii) provides uniqueness. As shown in Appendix, the same two conditions also ensure the existence and uniqueness of \( \Lambda \) as a function of quantities instead of normalized prices.

A caveat is that equation (27) may not necessarily admit a solution in \( U \), for a given \( q \) and \( \Lambda \). While there is no simple condition on \( D_i, F \) and \( H \) that would systematically ensure existence, in most practical examples it is easy to check that a solution exists. For instance, if neither \( F \), \( H \) and \( G \) depend on \( U \), as in several of the examples provided below in Section 4, a sufficient condition for the existence and monotonicity in \( U \) is that \( D_i(q_i, U) \) is strictly decreasing in \( U \) (holding \( q_i \) constant) and varies from \(+\infty\) to zero in the limit over the range of \( U \). Conversely, interesting cases also arise when only \( H \) and \( F \) depend on \( U \) (e.g. semi-separable preferences as discussed in Section 4.1).

Proposition 5 highlights how to characterize direct utility as a function of quantities \( q \). As in the single-aggregator case, we obtain a similar characterization of indirect utility as a function of normalized prices \( p/w \). Integrating by part, we show in the appendix that the indirect utility satisfies the following equation in \( V \):

\[
\sum_i \int_{y=y_0}^{\frac{p_i}{w}} F(\Lambda, V) D_i(y, V) dy = K(\Lambda, V) \quad (28)
\]

where \( K \) is such that \( \frac{\partial K}{\partial \Lambda}(\Lambda, V) = \frac{\partial F}{\partial \Lambda}(\Lambda, V)H(\Lambda, V) \), where \( \Lambda \) can again be implicitly defined such that the budget constraint holds (here as a function of normalized prices \( p/w \)). Using Roy’s identity, we can obtain Marshallian demand directly from this expression, which is sometimes simpler than expression (27) in Proposition 5 (e.g. as in cases of indirect separability).

In Section 4 below, we examine various examples where these results can be applied.

4 Special cases and examples

This section discusses additional examples and special cases, including a discussion of different forms of separability, several examples of homothetic preferences, and a practical illustration of how to add flexible income effects to a given set of demand curves in partial equilibrium. Here I also examine demand systems with two aggregators as in Thisse and Ushchev (2016), and show that one of the two aggregators can be set equal to indirect utility without loss of generalization. Finally, I discuss extensions to demand with choke prices.
4.1 Different forms of separability as special cases

Direct and indirect additive separability  Let us recall here the functional form taken in one of the most simple cases discussed earlier, direct additive separability, as it will serve as a reference for other generalizations. Preferences are directly-separable if there is only a single aggregator and function \( H \) is constant, and we can write utility as:

\[
U(q) = \sum_i \int_{q_i = q_{i0}}^{q_i} D_i^{-1}(q) dq
\]

which also leads to a simple demand function: \( q_i = D_i(\Lambda p_i / w) \). Directly-separable preferences have been discussed extensively in the literature, across many fields in economics. The main reason for their wide use is their tractability, and they already offer flexible price effects along each demand curve for each good. However, as pointed out for instance by Deaton (1974), assuming direct separability comes at the cost of imposing strong restrictions on price and income elasticities.

A first step away from directly-separable preferences is to consider indirectly-separable preferences, for which indirect utility can be written as

\[
V(p/w) = \sum_i \int_{y=y_{i0}}^{p_i/w} D_i(y) dy
\]

which leads to a demand function even more simple than in the previous case: \( q_i = D_i(p_i / w) / \Lambda \) with \( \Lambda = \sum_j (p_j / w)D_j(p_j / w) \). Bertoletti and Etro (2017) argue that these types of preferences also offer very tractable demand functions, and that they lead to better predictions than directly-separable preferences for instance if we examine and test how prices respond to market size and income in otherwise standard general-equilibrium models. However, these preferences still impose strong restrictions on demand patterns. For instance, as shown in equation (8), indirectly-additive preferences still impose a very tight link between income and price elasticities: \( \frac{\partial \log q_i}{\partial \log w} = 1 + \bar{\varepsilon}_D - \varepsilon_{Di} \) (as \( \varepsilon_H = 0 \) for indirectly-separable preferences), where \( \varepsilon_{Di} \) corresponds to the price elasticity of demand when good \( i \) has a negligible expenditure share.

The other examples provided below explore various ways to provide additional flexibility in price and income patterns while retaining simple functional form for practical applications.

Implicit additive separability  A type of separability which has recently seen a gain in interest is implicit (additive) separability, which again can be distinguished into direct and indirect implicit separability. Preferences are directly implicitly separable if utility can be
characterized as the solution of an equation of the type:\(^{19}\)

\[
\sum_i \int_{q=q_0}^{q_i} D_i^{-1}(q,U) dq = 1
\]  

(29)

where \(D_i\) is a function of two arguments.

Such preferences are a special case of Proposition 5 with two aggregators, but not Proposition 3. In fact, implicitly-additively separable preferences depend on a single aggregator only when they are also directly separable (when the price shifter \(F\) is constant) or when price elasticities are uniform (non-homothetic CES case). With two aggregators as in Proposition 5, we can show that preferences are implicitly additively separable if and only if \(F\) does not depend on \(\Lambda\), and in this case it is without loss of generality to assume \(F = 1\).

A similar result is obtained for the implicitly-indirectly-additive case, defined as when indirect utility can be characterized as the solution of:

\[
\sum_i \int_{y=y_0}^{p_i/w} D_i(y,V) dq = 1.
\]  

(30)

Demand as in Proposition 5 can be derived from indirectly-implicitly-additively-separable preferences only if \(H\) does not depend on aggregator \(\Lambda\). Conversely, demand derived from indirectly-implicitly-additively-separable preferences take the form described in Proposition 5 with \(H = 1\).

Implicit separability (direct or indirect) can prove useful in order to generate price and income effects that are less tightly related as with direct and indirect separability. In particular, for a given consumer, the ranking in price elasticities across goods can be totally uncorrelated with the ranking of income elasticities.\(^{20}\)

**Direct semi-separability** Let us introduce a class of preference which we could refer to as “semi-separable” where we can express either direct or indirect utility as a more simple function of quantities and prices as well as the aggregator. First, let us define preferences as directly semi-separable if we can write utility as:

\[
U(q) = \frac{1}{G(\Lambda)} \sum_i R_i(H(\Lambda)q_i)
\]  

(31)

where \(H\), \(G\) and \(R_i\) are twice continuously-differentiable, with \(G' > 0\), \(H' > 0\), \(R'_i > 0\) and \(R''_i < 0\). As with Gorman-Pollak demand, we define \(\Lambda\) such that the derivative w.r.t. \(\Lambda\) of the

---

\(^{19}\)Blackorby et al. (1991) provide yet another generalization of implicit separability.

\(^{20}\)Implicit separability, resp. direct and indirect, offers less flexibility than the general form of Proposition 5 for how demand can shift, resp. horizontally and vertically (with shifts that can only depend on utility).
expression above is null, i.e. such that:
\[
\frac{\sum_i q_i R_i'(H(\Lambda)q_i)}{\sum_i R_i(H(\Lambda)q_i)} = \frac{F(\Lambda)}{G(\Lambda)}
\] (32)

where \( F(\Lambda) \equiv G'(\Lambda)/H'(\Lambda) \) is assumed to be a positive and continuously differentiable.\(^{21}\)

This demand system is a special case of Proposition 5. This corresponds to defining \( D_i(y_i, V) = R_i^{-1}(V y_i) \), or equivalently: \( D_i^{-1}(q_i, U) = R_i'(U q_i)/U \), and specifying \( F, G, \) and \( H \) as functions of \( \Lambda \) only. Conditions [A5] required by Proposition 5 are met if \( R_i^{-1}(F(\Lambda)y_i)/H(\Lambda) \) has a strictly negative derivative in \( \Lambda \) and goes from +\( \infty \) to 0 (in the limit) as \( \Lambda \) increases, holding \( y_i \) fixed. In this case, the system of equations (31) and (32) has a unique solution in the aggregator \( \Lambda \) and utility \( U \), and define a well-behaved utility for any \( q \).

Demand for good \( i \) is then:
\[
q_i = \frac{R_i^{-1}(V F(\Lambda)p_i/w)}{H(\Lambda)}
\] (33)

where \( V = V(p/w) \) refers to indirect utility.

These preferences provide a generalization of directly-additive separability, and also retain some of the properties associated with direct separability. Directly-separable preferences correspond to the limit case where both \( H \) and \( G \) are constant and \( F(\Lambda) = \Lambda \). These preferences offer about the same degree of flexibility as Gorman-Pollak preferences with a single aggregator (the multiplicative specification of utility, equation (31) mirrors the additive specification in Proposition 3). For a given consumer, there is again an affine relationship between income elasticities and price elasticities across goods.

An advantage of semi-separability (relative to the more general specification in Proposition 5) is that direct utility can be expressed more simply and explicitly as a function of quantities and the aggregator \( \Lambda \). In general, there is no explicit expression for aggregator \( \Lambda \) and indirect utility \( V \), but this caveat also applies to directly-additive separability. Moreover, as will be discussed in Section 5.2, another reason to introduce this new type separability is to highlight a more general class of preferences with similar implications for market size effects as additively-separable preferences.

**Indirect semi-separability** We can obtain a similar functional form for indirect utility if we make the same functional form assumptions as above for \( D_i^{-1} \) instead of \( D_i \).

\(^{21}\)Since both \( G \) and \( H \) are strictly monotonic functions of \( \Lambda \), it is without loss of generality to impose either \( H = 1 \) or \( G = 1 \), whichever is more practical.
Suppose that indirect utility can be expressed as:

\[ V(p/w) = \frac{\sum_i S_i(F(\Lambda)p_i/w)}{L(\Lambda)} \]  

(34)

where \( F, L \) and \( S_i \) are twice continuously differentiable, with \( F' > 0, L' < 0, S_i' < 0 \) and \( S_i'' > 0 \). We define \( \Lambda \) such that the derivative w.r.t. \( \Lambda \) of the expression above is null, i.e. such that:

\[ \frac{\sum_i (p_i/w)D_i(F(\Lambda)p_i/w)}{\sum_i S_i(F(\Lambda)p_i/w)} = \frac{H(\Lambda)}{L(\Lambda)} \]  

(35)

where we denote \( D_i = -S_i' \) and define \( H(\Lambda) = -L'(\Lambda)/F'(\Lambda) \), a positive and continuously-differentiable function of \( \Lambda \). Note that this equation in \( \Lambda \) does not involve indirect utility \( V \).

Again, such indirect utility function is a special case of the dual-aggregator form that we studied in Proposition 5. This corresponds to defining \( D_i(y_i, V) = -S_i'(y_i)/V \) and specifying \( F \) and \( H \) as functions of \( \Lambda \) only. The conditions required by Proposition 5 are met if \( D_i(F(\Lambda)y_i)/H(\Lambda) \) has a strictly negative derivative in \( \Lambda \) and goes from \( +\infty \) to \( 0 \) (in the limit) as \( \Lambda \) increases. The system of equations (34) and (35) has a unique solution in the aggregator \( \Lambda \) and indirect utility \( V \), and correspond to a well-behaved utility for any \( q \). We also describe in Appendix how to characterize the direct utility function as in Proposition 5.

In this case, Marshallian demand takes the form:

\[ q_i = \frac{D_i(F(\Lambda)p_i/w)}{VH(\Lambda)} \]  

(36)

where \( V = V(p/w) \) is indirect utility and \( \Lambda \) can itself seen as a function of \( p/w \) using condition (35). As the name suggests, such preferences provide a generalization of indirectly-additive separability, which corresponds to the limit case where \( F \) and \( K \) are constant. Such preferences yield similar properties as indirectly-additively separable preferences in terms of market size effects in general-equilibrium models with economies of scale, as we will discuss in Section 5.

### 4.2 Homotheticity

There are many reasons for which one may want to impose homotheticity, e.g. to allow for simple aggregation properties across consumers with heterogeneous income levels, to provide a straightforward interpretation of price indices, or to model growth with multiple sectors under a balanced growth path. The homothetic two-aggregator specification described in this section offers a parsimonious yet flexible framework that encompasses various examples of homothetic preferences that have been used in the literature.
In the two-aggregator homothetic case, the demand shifters $F$ and $H$ can be expressed as a function of the aggregator $\Lambda$ only, while demand depend on both $\Lambda$ and the ideal price index $P$:

$$q_i = \frac{w}{H(\Lambda)P}D_i \left( \frac{F(\Lambda)p_i}{P} \right)$$

(37)

where aggregator $\Lambda$ can be implicitly defined by the budget constraint as in equation (26).

The ideal price index $P$ is then implicitly defined by the following equation:

$$\sum_i \int_{y=y_{i0}} \frac{p_i F(\Lambda)}{P} D_i(y)dy - \int_{\lambda=\Lambda_0} F'(\lambda)H(\lambda)d\lambda = c_0$$

(38)

Similarly, utility $U$ can be implicitly defined as the solution of:

$$\sum_i \int_{q=q_{i0}} \frac{q_i H(\Lambda)}{u} D_i^{-1}(q)dq - \int_{\lambda=\Lambda_0} H'(\lambda)F(\lambda)d\lambda = c_1$$

(39)

where $c_0$, $c_1$ and $\Lambda_0$ are constant terms. Note that $\Lambda$ is such that the partial derivative of the left-hand side w.r.t $\Lambda$ is null for both (38) and (39). It is also straightforward to check that the implicit solution for utility is homogeneous of degree one in quantities $q$.

In spite of imposing homotheticity, this specification offers rich price effects, especially if we compare them to CES preferences: it allows for a flexible specification of each demand curve thanks to $D_i$, and allows for competition (through the aggregator $\Lambda$) to shift demand curves vertically (through the price shifter $F$) or horizontally (through the quantity shifter $H$).

Four sub-cases are particularly interesting and correspond to the three cases studied by Matsuyama and Ushchev (2017), as well as Feenstra (2018)'s QMOR:

**Homothetic Single Aggregator.** This Gorman-Pollack demand system is homothetic if and only if $H(\Lambda)F(\Lambda)$ is constant or if it is CES.\(^{22}\) In this case, without loss of generality we can assume that $F(\Lambda) = 1/H(\Lambda) = \Lambda$. A homogeneous utility representation is given by:

$$\log U(q) = \log(\Lambda) + \sum_i \int_{x=x_{i0}}^{q_i/\Lambda} D_i^{-1}(x)dx$$

where $\Lambda$ is such that $\sum_i (q_i/\Lambda)D_i^{-1}(q_i/\Lambda) = 1$, and $x_{i0}$ are constant terms. In this case, the single aggregator $\Lambda$ is homogeneous of degree one in quantities $q_i$ in the primal version. We

\(^{22}\)More trivially, the non-homothetic CES case (case 2 of Proposition 1) is homothetic only in the standard homothetic CES case.
can also express $\Lambda$ as a function of prices $p_i$, and write expenditure shares as:

$$p_iq_i/w = \Lambda p_i D_i (\Lambda p_i).$$  

(40)

This specification is particularly attractive for empirical purposes, as it allows for flexible demand curves $D_i$ and yet a single aggregator $\Lambda$ to capture income as well as all other prices.

**Homothetic Direct Implicit Additivity.** When $H(\Lambda) = 1$ is constant, utility can be defined implicitly with a simple expression that does not involve aggregator $\Lambda$. In the homothetic case, this yields:

$$\sum_i \int_{q=q_0}^{q_i} D_i^{-1}(q) dq = 1.$$  

(41)

This case is described in Matsuyama and Ushchev (2017), and also corresponds to Kimball (1995) when $D_i^{-1}$ is identical across goods. Demand for good $i$ corresponds to:

$$q_i = (w/P)D_i (\Lambda p_i/P)$$  

(42)

where $\Lambda$ can again be defined implicitly by the budget constraint is satisfied.

**Homothetic Indirect Implicit Additivity.** Symmetrically, when $F(\Lambda) = 1$, indirect utility and the price index can be defined implicitly without involving aggregator $\Lambda$. For the ideal price index, we obtain:

$$\sum_i \int_{y=y_0}^{y_i} D_i(y)dy = 1.$$  

(43)

In this case, demand corresponds to:

$$q_i = wD_i (p_i/P) \sum_j p_j D_j (p_j/P).$$  

(44)

**Symmetric QMOR.** QMOR preferences have first been studied by Diewert (1976) and more recently studied by Feenstra (2018) imposing some symmetry in the price effects.\(^{23}\) Take $D_i(y) = \alpha_i y^{r-1} + \beta_i y^{\kappa r-1}$ and $F(\Lambda) = \Lambda$ and $H(\Lambda) = \Lambda^{-r-1}$ with $r < 0$ and $\kappa \in (0,1)$. In this case, we can obtain an explicit expression both for the price index and the aggregator $\Lambda$:

$$P^r = \sum_i \alpha_i p_i^r + \left( \sum_i \beta_i p_i^{\kappa r} \right)^{\frac{1}{\kappa}}; \quad \Lambda^{-\kappa r} = \sum_i \beta_i \left( \frac{p_i}{P} \right)^{\kappa r}.$$  

\(^{23}\)Here, I consider a slight generalization of demand system used in Feenstra (2018) by allowing $\kappa \neq 1/2$.  

26
Demand is then:

\[ q_i = \frac{w}{P} \left( \frac{p_i}{P} \right)^{r-1} \left[ \alpha_i + \beta_i \left( \frac{\Lambda p_i}{P} \right)^{-r(1-\kappa)} \right]. \]

With \( \kappa = 1/2 \), symmetric \( \alpha_i = \alpha \) and \( \beta_i = \beta \), we obtain the symmetric QMOR specification as in Feenstra (2018). When \( \alpha > 0 \) and \( \beta < 0 \), note that we get a finite reservation price (choke price). We discuss such possibility below in Section 4.5. In the next sub-section, we also discuss a non-homothetic generalization of symmetric QMOR.

Yet another example of homothetic demand with two aggregators is the homothetic translog cost functions and linear demand, discussed below in Section 4.5 and in Appendix. The specification described in (37) encompasses all these cases, and provides further generalizations.

### 4.3 Modeling richer income effects: a concrete example

Suppose that demand \( q_i \) for product \( i \) is provided by a demand curve \( \tilde{D}_i(p_i) \) in partial equilibrium, holding other prices constant (thus holding utility and other aggregates constant). Mrázová and Neary (2013) examine the shape and properties of these demand curves, as well as sufficient statistics to determine firms’ behavior in monopolistic competition models. There are many ways, however, to rationalize such demand curves. Mrázová and Neary (2013) indicate that any of such demand curve can be obtained from a directly-additive utility function, in which case demand can be fully specified as \( q_i = \tilde{D}_i(\Lambda p_i) \) where \( \Lambda \) captures the response to all other changes in prices and income.\(^{24}\) However, as discussed earlier, this leads to strong restrictions in terms of income effects.

By introducing one or two aggregators as in Propositions 3 and 5, one can instead rationalize such demand systems along with much richer Engel curves and income effects. First, we can derive such demand from a Gorman-Pollak demand system \( q_i = \tilde{D}_i(F(\Lambda)p_i)/H(\Lambda) \) where changes in other prices and income influence both the price shifter \( F \) and the quantity shifter \( H \). Going one step further, Proposition 5 shows that we can make such demand system even more flexible by specifying \( q_i = \tilde{D}_i(F(\Lambda,V)p_i,V)/H(\Lambda,V) \). We illustrate this approach more concretely below by examining bi-power demand curves.\(^{25}\)

**Bi-power demand.** A prominent type of demand studied in Mrázová and Neary (2013) is the bi-power form, where demand for good \( i \) takes the form:

\[ q_i = \gamma_i p_i^{\nu_i} + \delta_i p_i^{\sigma_i}, \]

\(^{24}\)and where \( \Lambda \) coincides with the Lagrange multiplier of the budget constraint.

\(^{25}\)See Appendix for examples based on conditionally-linear demand.
in partial equilibrium, i.e. holding other prices and income constant.\textsuperscript{26} This example is particularly relevant as it includes not only iso-elastic demand curves as special cases, but also a variety of other demand curves used in the literature, such as the PIGL family, the Pollak family, and QMOR.

Other prices and income may potentially affect all four determinants of the demand curve: $\gamma_i$, $\nu_i$, $\delta_i$ and $\sigma_i$. A property highlighted by Mrázová and Neary (2013) is that the relationship between the price elasticity and the curvature of demand (the "demand manifold") depends only on the exponents $\nu_i$ and $\sigma_i$, and is invariant to shocks in the demand shifters, $\gamma_i$ and $\delta_i$.

Allowing for a single price aggregator, Proposition 1 and 3 indicate that this aggregator $\Lambda$ can affect demand only through a common price shifter or quantity shifter. With a single aggregator, this implies that bi-power demand with a single aggregator must take the form:

$$q_i = \frac{\alpha_i [F(\Lambda) p_i / w]^{-\nu_i} + \beta_i [F(\Lambda) p_i / w]^{-\sigma_i}}{H(\Lambda)}$$

where $\alpha_i > 0$ and $\beta_i > 0$ are positive constant terms, and where $\Lambda$ adjusts to satisfy the budget constraint. Condition [A3]-ii) imposes $\min\{\nu_i \varepsilon_F, \sigma_i \varepsilon_F\} + \varepsilon_H > 0$ for any $\Lambda$. In particular, it may be convenient to specify iso-elastic demand shifters: $F(\Lambda) = \Lambda$ and $H(\Lambda) = \Lambda^{-\eta}$. Applying Proposition 3, such demand system can be rationalized as long as $\min\{\nu_i, \sigma_i\} > \eta$ or $\max\{\nu_i, \sigma_i\} < \eta$.\textsuperscript{27}

Allowing also for indirect utility as an additional aggregator, Proposition 2 indicate that bi-power demand must then take the form:

$$q_i = \frac{\alpha_i(V) [F(\Lambda, V) p_i / w]^{-\nu_i(V)} + \beta_i(V) [F(\Lambda, V) p_i / w]^{-\sigma_i(V)}}{H(\Lambda, V)}$$

where $\alpha_i$, $\beta_i$, $\nu_i$ and $\sigma_i$ are now functions of utility, and thus indirectly vary with income. As utility increases, different goods $i$ may be associated with smaller or larger demand, and may be associated with higher or smaller price elasticities.

It might again be convenient to restrict to iso-elastic demand shifters: $F(\Lambda) = \Lambda$ and $H(\Lambda) = \Lambda^{-\eta}$. For the demand manifold to remain invariant, we also assume that the exponents $\nu_i$ and $\sigma_i$ are constant, which yields:

$$q_i = \alpha_i(V) \Lambda^\eta [\Lambda p_i / w]^{-\nu_i} + \beta_i(V) \Lambda^\eta [\Lambda p_i / w]^{-\sigma_i}. \quad (45)$$

Applying Proposition 5, such demand system can be rationalized if $\min\{\nu_i, \sigma_i\} > \eta$ or $\max\{\nu_i, \sigma_i\} < \eta$.

\textsuperscript{26}We can also examine bi-power inverse demand in a similar fashion.

\textsuperscript{27}In the latter case, consider a change in variable $1/\Lambda$ to satisfy the conditions in Proposition 3.
\[ \eta \text{ and if the expression above is strictly decreasing in } V. \]\n
**A non-homothetic generalization of QMOR.** An interesting special case of (45) is when the coefficient \( \eta \) is equal to one of the two exponents for prices. This happens to provide a generalization of symmetric QMOR studied in Section (4.2). A convenient feature is that we can solve explicitly for the aggregator \( \Lambda \) as a function of indirect utility. Borrowing a similar functional form as homothetic QMOR, we can obtain a more general specification where price effects are very similar to QMOR, yet allow for more flexible Engel curves. Such generalization remains a special case of the two-aggregator demand systems described in Proposition 5.

Suppose that \( \nu_i = \nu > 1 \) and \( \sigma_i = \sigma > 1 \) are identical across all goods and that \( \sigma < \nu \), and suppose that \( \eta = \nu \), we can obtain an explicit solution for the aggregator \( \Lambda \) as a function of prices, utility and income:

\[
\Lambda^{\nu-1} = \sum_i \beta_i(V) \left( \frac{p_i}{w} \right)^{1-\sigma}.
\]

Indirect utility can then be seen as an implicit solution of an equation that no longer involve the aggregator \( \Lambda \):

\[
\sum_i \alpha_i(V) \left( \frac{p_i}{w} \right)^{1-\nu} + \left( \sum_i \beta_i(V) \left( \frac{p_i}{w} \right)^{1-\sigma} \right)^{\frac{1}{1-\nu}} = 1.
\]

If \( \alpha_i(V) \) and \( \beta_i(V) \) are positive, assuming that they strictly decrease with \( V \) provides a sufficient condition for this indirect utility function to coincide with rational consumer preferences.

Such demand system then yields a demand that features substitution and price effects that are very similar to homothetic QMOR:

\[
q_i = \alpha_i(V) \left( \frac{p_i}{w} \right)^{-\nu} + \beta_i(V) \left( \frac{p_i}{w} \right)^{-\sigma} \left( \sum_j \beta_j(V) \left( \frac{p_j}{w} \right)^{1-\sigma} \right)^{\frac{\sigma-\nu}{1-\sigma}}
\]

and now allows for richer income effects through the functions \( \alpha_i \) and \( \beta_i \) which can both flexibly demand on indirect utility. This demand system also provides a generalization of non-homothetic CES preferences described in Proposition 4 in the limit case where \( \nu = \sigma \). As noted previously, we could even allow \( \nu \) and \( \sigma \) to be functions of indirect utility \( V \), but the combination of \( \alpha_i(V) \) and \( \beta_i(V) \) already provide a way to parameterize how income affects the curvature of indifference curves.

---

\[\text{If } \alpha_i(V) \text{ and } \beta_i(V) \text{ are positive, a sufficient condition is that they both decrease with } V. \] We can also allow \( \beta_i(V) \) to be negative, which leads to choke prices as discussed in Section 4.5.
4.4 Double-shifter demand system

Thisse and Ushchev (2016) show that the following demand system can be obtained by aggregating over many consumers who make indivisible consumption choices among horizontally-differentiated product varieties:

\[ q_i = Q(p/w)D_i(F(p/w)p_i/w) \]  

where \( Q \) and \( F \) are two aggregators, i.e. two continuously-differentiable mappings from \( \mathbb{R}^N_+ \) to \( \mathbb{R}^N_+ \), and \( D_i \) is a continuously-differentiable mappings from \( \mathbb{R}_+ \) to \( \mathbb{R}_+ \). Note that the budget constraint implies:

\[ Q(p/w) = 1/\sum_j (p_j/w)D_j(F(p/w)p_j/w). \]

For instance, aggregate consumption that mimics indirectly-additive preferences can be obtained by aggregating over consumers with multinomial logit idiosyncratic utility terms. As discussed in Thisse and Ushchev (2016), more general aggregate consumption patterns can be obtained with alternative distributions of random utility terms across consumers.

The demand system specified in equation (46) is easy to manipulate, estimate, and provide a natural extension of directly-additive and indirectly-additive preferences which only include one of the two demand shifters. For instance, this specification of demand is used in Arkolakis et al. (2019) to obtain a so-called “gravity equation” for aggregate trade between countries, in a model with heterogeneous firms and asymmetric countries. Arkolakis et al. (2019) give examples of demand with such structure but do not provide a micro-foundation for the more general functional form.

At first sight, this does not appear to be a special case of the demand systems used in Proposition 2 and 5, which assume that the indirect utility \( V \) to be one of the two aggregators. However, with symmetry and rank arguments, we can show that the triplet of gradients of \( F, Q \) and \( V \) cannot have a rank higher than two if the demand system is integrable. Hence, such demand can actually be re-expressed using shifters that are functions of just utility \( V \) and another aggregator \( \Lambda \).

**Proposition 6** Suppose that the demand system takes the form given by (46), that it is integrable, and that the pair of gradients \[ \left\{ \frac{\partial Q}{\partial \log p}, \frac{\partial F}{\partial \log p} \right\} \] has rank two for all \((p, w)\). Demand can then be written as:

\[ q_i p_i/w = \tilde{Q}(\Lambda, V)D_i(\tilde{F}(\Lambda, V)p_i/w) \]  

for some functions \( \tilde{Q} \) and \( \tilde{F} \) of indirect utility \( V \) and a common aggregator \( \Lambda \).

Hence, when such demand is integrable (with a representative consumer) it can be seen as a special case of the demand systems examined in Proposition 5.
Such a demand system correspond to Gorman-Pollak demand with a single aggregator when $Q$ and $F$ can be written as a function of a single aggregator $\Lambda$ instead of two aggregators $(\Lambda, V)$, but in this case the gradients of the two aggregators, $\frac{\partial Q}{\partial \log p}$ and $\frac{\partial F}{\partial \log p}$, must be colinear.

### 4.5 Demand with choke prices

In various applications, demand from a consumer may be equal to zero if the price of a good is too high. Such upper bound is called a choke price or reservation price. This is an often desired feature for estimation, as zeros are prevalent in microdata at the finest level, and for applied modeling, e.g. to generate non-trivial extensive margins and explain selection across markets.\(^{29}\) One would want such choke price to be an equilibrium outcome, and depend on consumer income and the toughness of competition. For instance, there is substantial evidence of income effects, e.g. from Hummels and Klenow (2005): richer consumers buy a wider range of products and richer countries import a larger variety of goods.

The Gorman-Pollak and dual-aggregator demand system studied above can be accommodated to yield such choke prices. With a demand structure as in Proposition 5, suppose that $D_i(y_i, V) = 0$ for all $y_i \geq a_i(V)$ in the dual-aggregator case — this becomes $D_i(y_i) = 0$ for all $y_i \geq a_i$ in the single-aggregator case (Gorman-Pollak demand, Proposition 3). Most of the results shown previously hold if, with a slight abuse of notation, we define $D_i^{-1}(0, V) = a_i(V)$.

In this framework, the choke price $p_i^*$ depends on income, utility and the aggregator $\Lambda$. For a consumer with income $w$, aggregator $\Lambda$ and utility $V$, demand for good $i$ is null if and only if:

$$p_i \geq p_i^* = \frac{a_i(V)w}{F(\Lambda, V)}.$$

With a single aggregator, the choke price has a more restrictive functional form: $p_i^* = \frac{a_iw}{F(\Lambda)}$.

The choke price is proportional to income when preferences are indirectly additively separable since the terms $a_i$ and $F$ are constant in this case. Bertoletti et al. (2018) exploit this property to obtain a tractable model of trade and argue that it fits key patterns of how prices vary with income and population across markets.\(^{30}\) A similar property can be obtained with indirectly implicitly separable preferences (see Section 4.1) as the choke price would then just depend on income and utility.

A very simple and tractable case of demand with choke prices is demand that is linear in

\(^{29}\)Choke prices are particularly useful in international trade to explain why less efficient firms are less likely to export to a specific market (without having to rely on export fixed costs) and to obtain gravity equations as shown in Melitz and Ottaviano (2008) and Arkolakis et al. (2019) among others.

\(^{30}\)Within a standard model of trade with choke prices, Fally (2019) examines an application with a single aggregator and iso-elastic demand shifters $F$ and $H$ in order to generalize some insights by Arkolakis et al. (2019) and Bertoletti et al. (2018) on how the gains from trade depend on the structure of preferences.
its own price. As described in Appendix, there is a variety of ways to generate such demand by allowing for one or two aggregators that influence how other prices and income shift the demand curve vertically and horizontally.

Another tractable example of preferences used in the macroeconomic and trade literature is the Translog expenditure function (Feenstra, 2003; Bergin and Feenstra, 2009; Novy, 2013; Feenstra and Weinstein, 2017), which generates choke prices. A typical assumption is that the cross-price elasticities are symmetrical. Demand associated with Translog can then be expressed as a function of a single aggregator \( \Lambda \) even when some varieties are not consumed (see Appendix), with expenditure shares taking the form:

\[
p_i q_i / w = \alpha_i - \gamma \log(\Lambda p_i / w)
\]

with a choke price \( p^*_i = \exp(\alpha_i / \gamma_i) w / \Lambda \).

Yet another example of preferences with two aggregators generating choke prices is QMOR, as well as its non-homothetic extension, with choke prices arising when \( \beta_i(V) \) is negative.

5 An application to monopolistic competition

Summarizing other prices by a one (or two) aggregator is particularly useful for applications to imperfect competition, as such aggregator synthetizes all relevant information on a firm’s competitors. Under monopolistic competition, assuming that each firm has a negligible market share (as in Dixit and Stiglitz, 1977), this aggregator can be taken as given by a specific firm.\(^{31}\) This facilitates theoretical analysis of equilibrium as well as empirical estimation, while allowing for flexible equilibrium outcomes and comparative statics. This section discusses additional restrictions needed for such applications with a continuum of goods, then examines a simple general-equilibrium model with free entry under monopolistic competition to illustrate the role of modeling choices on the demand side.

5.1 With a continuum of goods

Models of monopolistic competition typically assume a continuum of product varieties\(^{32}\), where each variety accounts for a measure zero of aggregate expenditures. Here we discuss additional

\(^{31}\)The tools developed by Anderson et al. (2018) could be used in this case, using \( \Lambda \) as an aggregate. Under Bertrand competition, a firm with non-negligible market share would account for the effect of its own price on \( \Lambda \), holding other prices as given. Under Cournot competition, a firm would account for the effect of its own production quantity on \( \Lambda \), holding other quantities as given, using the inverse demand formulation and specifying \( \Lambda \) as a function of quantities instead of prices.

\(^{32}\)See e.g. Romer (1990), Grossman and Helpman (1991), Melitz (2003), Zhelobodko et al. (2012).
assumptions that should be imposed on the structure of demand such that it is well behaved on a continuum.

The discussion provided here fits within the framework of Parenti et al. (2017). A first assumption is that the set of potential varieties is compact and is included in \([0, \bar{N}]\); such assumption is typically not restrictive and this upper bound \(\bar{N}\) not binding in equilibrium if there is a fixed cost of producing a new variety and if \(\bar{N}\) is large enough. A consumption profile \(q\) is now defined as a mapping from \([0, \bar{N}]\) to \(\mathbb{R}_{\geq 0}\) that belongs to \(L^2([0, \bar{N}])\), i.e. such that its square has a finite integral sum.\(^{33}\) In this framework, utility and the aggregator \(\Lambda\) are two functionals, i.e. real valued functions defined over \(L^2([0, \bar{N}])\). They are assumed to be symmetric over \([0, \bar{N}]\), i.e. that consumers are indifferent to switching labels across products \(i\); here, this implies that function \(D_i = D\) is identical across all goods \(i\).

While strict quasi-concavity implies that consumers exhibit love for variety, we need to assume that utility does not drop too much when the quantity consumed \(q_i = 0\) is zero for a non-trivial measure of goods. To be more precise, here we assume \(\int_0^a D^{-1}(x)dx < \infty\) (a finite integral sum around zero). This implies that the expenditure share on a range of goods is zero in the limit if the quantity for these goods goes to zero (i.e. no good is essential): \(\lim_{q_i \to 0+} q_i D^{-1}(q_i H(\Lambda, U), U) = 0\). A sufficient condition for these properties to hold is that the elasticity of \(D\) is strictly larger than unity (or infinite) in the limit where the quantity of a good goes to zero.

Extending Proposition 5 to a continuum, utility \(U(q)\) needs to satisfy:

\[
\int_0^{\bar{N}} \int_{q=0}^{q_i H(\Lambda, U)} D^{-1}(q, U) \, dq \, di - G(\Lambda, U) = 0
\]

(48)

where aggregator \(\Lambda\) is itself an implicit solution to:

\[
\int_0^{\bar{N}} q_i D^{-1}(q_i H(\Lambda, U), U) di = F(\Lambda, U)
\]

(49)

and where \(D^{-1}\), \(H\), \(F\) and \(G\) are continuously differentiable real functions. Uniqueness is ensured by assuming that \(\varepsilon_D < \varepsilon_H\) and that the left-hand-side of (48) has a negative partial derivative in \(U\). A sufficient condition for existence of \(\Lambda\) (conditional on \(U\)) is that \(\frac{D^{-1}(q_i H(\Lambda, U), U)}{F(\Lambda, U)}\) takes on values from \(+\infty\) to 0 over the range of \(\Lambda\). Existence of utility is then guaranteed if we combine the following two conditions: i) we assume that \(\int_{q=0}^{H(\Lambda, U)} D^{-1}(q, U) / G(\Lambda, U) \, dq\) spans

\(^{33}\)\(L^2([0, \bar{N}])\) is a natural space on which to define consumption profiles as it is a Hilbert space and includes all bounded consumption profiles. Parenti et al. (2017) use this property to prove existence of an equilibrium. A less elegant alternative to obtain completeness would be to assume an uniform upper bound on the consumption profile \(q\) within a consumer’s budget set if such upper bound is not binding in equilibrium.
from $+\infty$ to 0 as utility decreases (holding $\Lambda$ and $q_i$ constant); ii) we assume that it goes to zero as $\Lambda$ tends to zero, for a any given $U$ and $q_i$.

Finally, a key assumption imposed by Parenti et al. (2017) is that utility is Frechet-differentiable in any $q \in L^2[0, N]$, which provides a rigorous definition of marginal utility in this context with a continuum of goods. Conditions to ensure Frechet-differentiability of $U$ and $\Lambda$ are discussed in Appendix.\footnote{Here one might be able to relax the requirement of Frechet differentiability given the existence of one or two aggregators summarizing all cross-price effects.}

While we focus here on symmetric demand, we refer to Bertoletti and Etro (2018a) for a discussion of the assumptions and approximations required under monopolistic competition when preferences are asymmetric across product varieties.

5.2 Market size effects

To illustrate the role of the demand side and in particular how assumptions and modeling choices influence key outcomes, the remainder of this section examines a simple general-equilibrium model with free entry under monopolistic competition with homogeneous firms. In particular, the goal is to examine how changes in market size (either from changes in population or productivity) affects firm size, prices and the number of firms depending on functional form assumptions on the demand side. A more elaborate study with heterogeneous firms, several markets and richer interactions is however beyond the scope of the present paper.

Model setup. Consider a single economy with a population $L$ of identical consumers. There is a continuum of products, each of them produced by a single firm, where $N$ denotes the measure of active firms. There is free entry of firms, who compete under monopolistic competition. Consumer preferences are described by those in the previous sub-section, with utility $U$ and aggregator $\Lambda$ satisfying equations (48) and (49).

There is only one factor of production, labor. We assume that $w$ is the efficiency of each worker, with $L$ is the number of workers, so that $Lw$ is the supply of labor in efficiency units. We normalize the return of a unit of labor to unity, which implies that the income of each worker (and consumer) equals $w$, and total GDP is given by $Lw$.

All firms have access to the same technology and cost structure, so firms are homogeneous. $Q$ denotes total production by firm, while $Q/L$ is the quantity consumed by variety and by worker. For each firm, the cost of producing $Q$ is given by a constant marginal cost $c$ and a fixed cost $f$, so that total costs equal $C(Q) = cQ + f$ in terms of efficiency units of labor. With a continuum of firms under monopolistic competition, each firm takes aggregates as given and
unaffected by its decisions, including utility $U$ and the aggregator $\Lambda$.

**Equilibrium conditions.** Two equilibrium conditions describe the supply side.

First, firms maximize profits. Sales for each firm are equal to production $Q$ times the price $p = wD^{-1}(HQ/L, U)/F$ where $F$ and $H$ themselves depend on aggregator $\Lambda$ and utility $U$. Profits are thus: $\pi = \max_Q \{ QwD^{-1}(HQ/L, U)/F - cQ - f \}$. Maximizing over $Q$ (taking $\Lambda$ and $U$ constant under monopolistic competition) leads to the usual first order condition equating markups and the inverse of the price elasticity of demand:

$$\frac{p - c}{p} = \frac{- (HQ/L)(D^{-1})'(HQ/L, U)}{\frac{D^{-1}(HQ/L, U)}{1/\sigma(HQ/L, U)}}$$

with $p/w = D^{-1}(HQ/L, U)/F$. The right-hand side is the inverse of the price elasticity of demand, $\sigma(HQ/L, U)$, which can be expressed as a function of utility $U$ as well as consumption quantity $Q/L$ multiplied by the quantity shifter $H(\Lambda, U)$.\(^{35}\)

Next, free entry implies that firms make zero profits in equilibrium: $\pi = 0$. Rearranging, this leads to the price $p$ equal to the average cost for each firm:

$$p = wD^{-1}(HQ/L, U)/F = (cQ + f)/Q.$$  

Two equilibrium conditions describe the demand side: equations (48) and (49) described above. With symmetry across product varieties, utility $U$ is such that:

$$N \int_{q=0}^{H(\Lambda, U)Q/L} D^{-1}(q, U) dq = G(\Lambda, U)$$

while the budget constraint can be written:

$$(NQ/L) D^{-1}(H(\Lambda, U)Q/L, U) = F(\Lambda, U).$$

We define an equilibrium as a set of $(Q, N, U, \Lambda)$ satisfying conditions (50), (51), (52) and (53).

**Market size effects across preferences specifications.** A central question, with implications for various fields in economics, is how prices and firm size depend on market size, where market size itself can be thought as the product of population and per capita income. As shown e.g. in Parenti et al. (2017), how we specify preferences on the demand side has sharp

\(^{35}\)The second order condition requires that the elasticity of $(D^{-1})'$ is larger than $-2$, which is assumed.
implications for comparative statics. In particular, in such a model, price $p$ and firm size $Q$ is independent of income $w$ when preferences are directly additively separable; independent of population $L$ when preferences are indirectly additively separable; and fully determined by total GDP when preferences are homothetic.

Added flexibility on the structure of the demand side can thus lead to a wider range of outcomes and comparative statics relative to imposing specific forms of separability. Here we use this framework to ask two questions. First, what are the comparative statics with a milder form of separability, e.g. with a single aggregator? Second, is the independence in income $w$ or population $L$ specific to directly-additive and indirectly-additive separability?

Demand with a single aggregator already generalizes both directly-additive and indirectly-additive preferences, hence it is not surprising that it encompasses a wide range of cases and comparative statics. Comparative statics depend on whether the demand shifters $F$ and $H$ depend on the aggregator $\Lambda$, recalling that $H$ is constant with directly-additive separability and $F$ is constant with indirectly-additive separability. Moreover, comparative statics depend crucially on whether demand is “superconvex” or “subconvex” (see e.g. Mrázová and Neary, 2013). We say that demand is superconvex if the price elasticity of demand $\sigma$ increases with sales, holding aggregator $\Lambda$ constant (i.e. if $\sigma'$ is monotonically increasing), and subconvex otherwise. As earlier, $\varepsilon_F$ and $\varepsilon_H$ denote the elasticity of $F$ and $H$ with respect to $\Lambda$.

**Proposition 7** With Gorman-Pollak demand with a single aggregator (as in Proposition 3):

i) An increase in population $L$ leads to an increase in firm size and a decrease in prices iff $\varepsilon_F < 0$ in the superconvex case ($\sigma' > 0$) or $\varepsilon_F > 0$ in the subconvex case ($\sigma' < 0$).

ii) An increase in income $w$ leads to an increase in firm size and a decrease in prices iff $\varepsilon_H > 0$ in the superconvex case ($\sigma' > 0$) or $\varepsilon_H < 0$ in the subconvex case ($\sigma' < 0$).

For instance, using directly-separable preferences ($\varepsilon_H = 0$ and $\varepsilon_F > 0$), Krugman (1979) assumes subconvexity to ensure that markups decrease with market size $L$. But none of these cases from Proposition 7 are ruled out by the assumptions needed for rationalization. As discussed along with Proposition 3, we assume $\varepsilon_D < 0$ and $\varepsilon_D\varepsilon_F < \varepsilon_H$, but we do not impose either super or sub-convexity, nor do we impose the sign of $\varepsilon_F$ and $\varepsilon_H$.

Here, we focus on comparative statics for firm size and prices, but similarly flexible outcomes can be obtained for the number of firms $N$. In a more detailed analysis, Bertoletti and Etro (2018b) also study market size effects within a similar model and demand structure, imposing either $\varepsilon_H = 1$ or $\varepsilon_F = 1$; in addition, they examine deviations from first-best allocations and
extensions with heterogeneous firms. Along the same lines, Fally (2019) shows that such single-aggregator demand systems can be used to obtain a broader range of predictions for the gains from trade, conditional on observed import penetration and elasticity of trade to trade costs.

A variety of comparative statics can also be achieved for instance with (directly or indirectly) implicitly-additive preferences. Based on the specification of equation (29), the price elasticity of substitution can be a flexible function of both quantities \( Q/L \) and the level of utility \( U \), which itself depends on the number of firms \( N \). As shown by Parenti et al. (2017), flexibility with respect to these two arguments allows generating a wide gallery of comparative statics (see Appendix). In particular, using implicit CES as in Proposition 4, an increase in income can lead to either an increase or decrease in the price elasticity \( \sigma(U) \), and thus either an increase or decrease in firm size in equilibrium.

Conversely, another question we ask in this framework is whether the insensitivity of production and prices to income is specific to directly-additive separable preferences and whether the insensitivity of production and prices to population is specific to indirectly-additive separable preferences. The answer is no, as we can extend these results to semi-separable preferences described in Section 4.1.

**Proposition 8** Suppose that preferences are semi-separable, as defined in Section 4.1:

i) If preferences are directly semi-separable, production \( Q \) and price \( p \) depend on population \( L \) but not on income \( w \), while the number of firms \( N \) is proportional to income \( w \).

ii) If preferences are indirectly semi-separable, production \( Q \) and price \( p \) depend on income \( w \) but not on population \( L \), while the number of firms \( N \) is proportional to population \( L \).

To show the first part i) on directly semi-separable preferences, it is useful to combine the free entry condition (51) with the pricing condition (50) and obtain a condition that only depends on production \( Q \) and the aggregator \( \Lambda \) (when function \( H \) does not directly depend on utility):

\[
- \frac{(HQ/L)(D^{-1})'(HQ/L)}{D^{-1}(HQ/L)} = \frac{f}{cQ + f}
\]  

As was described for instance in Zhelobodko et al. (2012), this states that under monopolistic competition we must have equality between the elasticity of revenues and the elasticity of costs \( \text{w.r.t.} \) firm production, \( Q \). On the demand side, direct semi-separability implies condition (32) which can now be rewritten as function of just consumption \( Q/L \) and the aggregator \( \Lambda \) (as \( H \) is itself a function of \( \Lambda \)):

\[
\frac{(HQ/L)}{\int_{q=0}^{HQ/L} D^{-1}(q) dq} = \frac{G'()}{G()} \]  

(55)
Combined, equations (54) and (55) determine $Q$ and $\Lambda$, and depend on population $L$ but not on income $w$. Hence production $Q$ and the aggregator $\Lambda$ (as well as the shifter $H$) are independent of income $w$ when preferences are directly semi-separable. Consequently, the average cost and thus the price $p$ are also independent of $w$. In turn, if firm size does not depend on $w$, the number of firms $N$ must be proportional to $w$.

A similar result is obtained in part ii) for indirect semi-separability. On the supply side, an expression analogous to (54) provides markups as a function of $F(\Lambda)$ and the normalized price $p/w$:

$$
\frac{p - c}{p} = \left[ \frac{(pF/w)D'(pF/w)}{D(pF/w)} \right]^{-1}
$$

On the demand side, with symmetric prices and demand across goods, condition (35) yields:

$$
\frac{(pF/w)D(pF/w)}{\int_0^{pF/w} D(y) dq} = \frac{HF}{S}
$$

(56)

These two equations jointly characterize the price $p$ and the aggregator $\Lambda$, as a function of income $w$ but not population $L$, hence neither $p$ and $\Lambda$ depend on population in equilibrium. As the price is equal to the average cost, $p = (cQ + f)/Q$, we also obtain that firm size, $Q$, does not depend on population and only varies with consumer income $w$.

Finally, another type of independence is obtained for homothetic preferences. In this case, the results of Parenti et al. (2017) also apply: production $Q$, the number of firms $N$ and prices $p$ (relative to the unit cost of labor) only depend on total GDP (i.e. $Lw$) but do not depend on $L$ and $w$ individually, conditional on total GDP.

Taken together, Propositions 7 and 8 illustrate how functional form assumptions made on the demand side influence key results on market size effects related to firm size, entry and prices in general equilibrium models. These results highlight both the need for flexible forms unless we want to purposefully shut down some specific channels, and highlight how demand with one or two aggregators ($\Lambda$ and $U$) can provide a rich and tractable framework.

6 Concluding remarks

Economists have often focused on demand systems where prices are conveniently summarized by a single aggregator, and where demand depends solely on such an aggregator, total expenditures and a good’s own price (“generalized separability”, following the terminology of Pollak 1972).

\footnote{Note also equation (55) does not depend on the number of product varieties $N$, hence the price elasticity of demand does not depend on $N$. Parenti et al. (2017) have shown that this implies that income does not affect firm size and prices in this framework.}
Here I show that such a demand system can take only one of two forms when price effects are not trivial. This result was already known by Pollak (1972) and Gorman (1972) but not formally demonstrated and is not well known today in spite of its usefulness. Furthermore, I show that these two types of demand systems can be rationalized (i.e. can be derived from well-behaved utility functions) under fairly mild regularity restrictions that guarantee a well-behaved quasi-concave utility.

The first case of demand allows for flexible price effects but more restricted income effects. The second case of demand allows for flexible income effects (Engel curves) but more restricted price effects. Allen-Uzawa substitution elasticities have to be constant across goods to ensure the symmetry of the Slutsky matrix but they may vary (increase or decrease) with utility and thus vary indirectly with income.

I further extend these results to demand systems that allow for two aggregators, one being the indirect utility function. I characterize the functional form that such demand must take, provide sufficient conditions to ensure that it can be rationalized, and characterize the utility function associated with such demand systems. Allowing for two aggregators can be useful for at least two reasons. First, it jointly allows for more flexible price and income effects than in the single-aggregator case, since the latter impose restrictions on either price or income effects. Second, it encompasses various examples of demand systems used frequently in the literature, thus providing a unified general structure.

Special cases and examples discussed here include: directly and indirectly additively-separable preferences, directly and indirectly implicitly-separable preferences, a new type of separability dubbed “semi-separability”, the three types of homothetic demand described in Matsuyama and Ushchev (2017), QMOR preferences as in Feenstra (2018), and dual-aggregator demand as in Thisse and Ushchev (2016) and Arkolakis et al. (2019).

There can be numerous applications and uses of such demand systems with one or two aggregators. Recent research in macroeconomics, international trade, industrial organization and development economics have highlighted in different contexts the crucial role of the demand side and its interactions with income disparities, fostered by an increased availability of precise micro-data on consumption baskets across households, such as scanner data. This paper aims to provide useful tools to model richer price and income effects in a tractable manner, for both theoretical and empirical applications.

References

Antonelli, G. B. (1886). *Sulla teoria matematica della economia politica*.


Appendix

Proofs of propositions and additional derivations

Proposition 1

Preliminaries: Inverse demand

Consider the demand system:

\[ q_i = \tilde{q}_i(p_i/w, \Lambda(p/w)). \]

Following condition [A1]-iv), for the sake of exposition we assume for the most part that for any \( q \in \mathbb{R}^N_+ \), there exists a vector of normalized prices \( p/w \in \mathbb{R}^N_+ \) that generates demand \( q \), i.e. such that \( q_i = \tilde{q}_i(p_i/w, \Lambda(p/w)) \).\(^{37}\)

First, note that \( \Lambda \) can be seen an implicit function of normalized prices \( p_i/w \) such that the budget constraint holds, i.e. such that:

\[ \sum_i (p_i/w) \tilde{q}_i(p_i/w, \Lambda) = 1. \]

If we assume that each \( q_i(p_i/w, \Lambda) \) is strictly decreasing in \( \Lambda \) (here we assume a strictly negative derivative), the solution in \( \Lambda \) is unique and continuously differentiable.

Since we assume that expenditure shares \( (p_i/w) q_i(p_i/w, \Lambda) \) monotonically decreases or increases with prices (holding \( \Lambda \) constant), demand can be inverted such that expenditure shares can be obtained as a function of \( q_i \) and the aggregator \( \Lambda \):

\[ q_i p_i/w = W_i(q_i, \Lambda) \]

i.e. such that \( (p_i/w) \tilde{q}_i(p_i/w, \Lambda) = W_i(\tilde{q}_i(p_i/w, \Lambda), \Lambda) \) for any \( \Lambda = \Lambda(p/w) \) and \( p/w \in \mathbb{R}^N_+ \). As demand \( \tilde{q}_i(p_i/w, \Lambda) \) has a strictly negative derivative in \( \Lambda \) (by assumption), by the implicit theorem we can also conclude that \( W_i \) has a strictly negative derivative in \( \Lambda \). Then we can also redefine \( \Lambda \) as an implicit differentiable function \( \Lambda(q) \) of the vector of quantities such that the budget constraint holds, i.e. such that: \( \sum_i W_i(q_i, \Lambda) = 1. \) As an abuse of notation, \( \Lambda \) denotes the aggregator both as a function of normalized prices and as a function of quantities \( q \) given that they coincide when \( q \) is the demand associated with normalized prices \( p/w).^{38}\)

In the remainder of the proof, since we focus on inverse demand, \( \Lambda \) primarily refers to such a function of quantities \( q \) rather than normalized prices \( p/w \).

Proof of Proposition 1

As described just above, the proof of Proposition 1 relies on the inverse demand function (using expenditures shares \( W_i(q, \Lambda) \) as functions of quantities and the aggregator \( \Lambda \)) rather than direct demand, and \( \Lambda \) is defined as a function of quantities \( q \), where \( W_i(q_i, \Lambda) \) is twice differentiable with a negative derivative in \( \Lambda \) and non-zero derivative in \( q_i \).

\(^{37}\)Alternatively, Proposition 1 applies to the image \( Q = \{ q : q_i = \tilde{q}_i(p_i/w, \Lambda(p/w)) > 0, p/w \in \mathbb{R}^N_+ \} \subset \mathbb{R}^N_+ \).

\(^{38}\)As a side note, we can also show that iso-\( \Lambda \) curves are connected, which implies that if any differentiable function of \( q \) that have a gradient that is proportional to the gradient of \( \Lambda \) (w.r.t. \( q \)) can be expressed as a function of \( \Lambda \).
Differentiating the budget constraint $\sum_i W_i(q_i, \Lambda) = 1$ w.r.t. $q_i$ implies:

$$\varepsilon_j(q_j, \Lambda) = \frac{S(q)}{W_j} \frac{\partial \Lambda}{\partial \log q_j}$$

(57)

where $\varepsilon_j(q_j, \Lambda) \equiv \left. \frac{\partial \log W_j}{\partial \log q_j} \right|_\Lambda$ denotes the elasticity w.r.t. own quantity $q_j$, holding aggregators constant, and where $S(q) \equiv \sum_i \frac{\partial W_i(q, \Lambda)}{\partial q_i}$ is strictly negative.

For such a demand system to be integrable and satisfy Slutsky symmetry, there must exist a utility function $U(q)$ and another real function $\lambda$ such that $\lambda(q) > 0$ and:

$$\frac{\partial U}{\partial \log q_i} = \lambda(q) W_i(q_i, \Lambda)$$

for any $q$. As mentioned in the text, we further assume that $U$ is twice continuously differentiable. Differentiating again, we obtain:

$$\frac{\partial U}{\partial \log q_i \partial \log q_j} = \frac{\partial \lambda}{\partial \log q_j} W_i + \lambda \frac{\partial W_i}{\partial \Lambda} \frac{\partial \Lambda}{\partial \log q_j}.$$  

The existence and continuity of the derivatives imply that the cross derivative is symmetric, hence:

$$\left( \frac{1}{W_j} \frac{\partial \log \lambda}{\partial \log q_j} \right) + \frac{\partial \log W_i}{\partial \Lambda} \left( \frac{1}{W_j} \frac{\partial \Lambda}{\partial \log q_j} \right) = \left( \frac{1}{W_i} \frac{\partial \log \lambda}{\partial \log q_i} \right) + \frac{\partial \log W_j}{\partial \Lambda} \left( \frac{1}{W_i} \frac{\partial \Lambda}{\partial \log q_i} \right).$$

Incorporating the expression from 57, this is equivalent to:

$$\left( \frac{S}{W_j} \frac{\partial \log \lambda}{\partial \log q_j} \right) + \frac{\partial \log W_i}{\partial \Lambda} \varepsilon_j = \left( \frac{S}{W_i} \frac{\partial \log \lambda}{\partial \log q_i} \right) + \frac{\partial \log W_j}{\partial \Lambda} \varepsilon_i.$$

holds for any $i \neq j$. Define $A_i(q) = \frac{S(q)}{W_j} \frac{\partial \log \lambda}{\partial \log q_j}(q)$, we obtain a key symmetry requirement that we will exploit below:

$$A_j(q) + \frac{\partial \log W_i}{\partial \Lambda}(q_i, \Lambda) \varepsilon_j(q_j, \Lambda) = A_i(q) + \frac{\partial \log W_j}{\partial \Lambda}(q_j, \Lambda) \varepsilon_i(q_i, \Lambda).$$

(58)

Next, we can see that we will be in either of these three cases (almost everywhere) in a neighborhood of any $q$:

- $Q_1$ is the set of vectors of quantities $q$ such that $\varepsilon_i(q_i, \Lambda)$ takes at least two different values across goods $i$ *even if we exclude any one good*.

- $Q_2$ is the set of vectors of quantities $q$ such that $\varepsilon_i(q_i, \Lambda)$ are identical across goods $i$.

- $Q_3$ is the set of vectors of quantities $q$ such that all $\varepsilon_i(q_i, \Lambda)$ are identical for all but one good.

For a neighborhood around $q$, suppose that $\frac{\partial \varepsilon_i}{\partial q_i}(q_i, \Lambda) \bigg|_\Lambda \neq 0$ and $\frac{\partial \varepsilon_j}{\partial q_j}(q_j, \Lambda) \bigg|_\Lambda \neq 0$ for at least two goods $i$ and $j$. In that case, we can see that $\varepsilon_i$, $\varepsilon_j$ and $\varepsilon_k$ will differ almost everywhere in a neighborhood of $q$ for $i$, $j$ and any third good $k$; hence we are in $Q_1$ almost everywhere around $q$. This is the first case considered below.
Next, suppose that $\left. \frac{\partial \varepsilon_i}{\partial q_{i0}} \right|_{\Lambda} \neq 0$ for just one good $i0$, i.e. $\varepsilon_j(q_j, \Lambda)$ does not depend on $q_j$ for goods other than $i0$ in the neighborhood of $q$. If $\varepsilon_j(q_j, \Lambda)$ takes two different values across goods $j$, we are in case 1. If $\varepsilon_j(q_j, \Lambda)$ is identical across all goods $j \neq i0$, we are in case 3 below.

Finally, if $\left. \frac{\partial \varepsilon_i}{\partial q_{i0}} \right|_{\Lambda} = 0$ in a neighborhood of $q$, we are either in case 1 below (if $\varepsilon_i$ takes on at least two different values even if we exclude a single good), in case 2 (if $\varepsilon_i$ is identical across all goods), or in case 3 (if $\varepsilon_i$ is identical across all but one good).

**Case 1** In an open set of $q$, suppose that $\varepsilon_i(q_i, \Lambda)$ takes at least two different values across goods $i$, even if we exclude any one good.

In this case, even if we exclude a single good $j$, there exists a vector $x_i(q)$ such that $\sum_i x_i = 0$ and $\sum_i \varepsilon_i x_i \neq 0$. Multiplying Equation (58) by $x_i(q)$ and summing up across goods $i$ (for a given $j$), we obtain:

$$
\left( \sum_i x_i \frac{\partial \log W_i}{\partial \Lambda} \right) \varepsilon_j = \left( \sum_i x_i A_i \right) + \left( \sum_i x_i \varepsilon_i \right) \frac{\partial \log W_j}{\partial \Lambda}.
$$

As $\sum_i \varepsilon_i x_i \neq 0$, we obtain that there exists two functions $h(q)$ and $m(q)$ such that:

$$
\frac{\partial \log W_j}{\partial \Lambda} = h(q) \varepsilon_j(q_j, \Lambda) + m(q).
$$

In particular, this holds also for any pair of goods $i$ and $j$. Taking the difference, we get:

$$
\frac{\partial \log W_j}{\partial \Lambda} - \frac{\partial \log W_i}{\partial \Lambda} = h(q) \left( \varepsilon_j(q_j, \Lambda) - \varepsilon_i(q_i, \Lambda) \right).
$$

In particular, take two goods for which $\varepsilon_i \neq \varepsilon_j$. Note that the left-hand side only depends on $q_j$, $q_i$ and $\Lambda$. This implies that $h(q)$ can be written as a function of $q_j$, $q_i$ and $\Lambda$ only.

If we’re not in case 3, we can also find a third good $i'$ such that $\varepsilon_{i'} \neq \varepsilon_i$ and $\varepsilon_{i'} \neq \varepsilon_j$. Applying the same argument, it must be that $h$ can be written as just a function of $\Lambda$, so we now denote it as $h = h(\Lambda)$.

Taking again a derivative in log $q_j$, holding $\Lambda$ constant, and noticing that the cross derivative is symmetric, $\frac{\partial \varepsilon_j}{\partial \Lambda} = \frac{\partial \log W_j}{\partial \log q_j} = \frac{\partial \log W_j}{\partial \log q_j}$, we obtain:

$$
\frac{\partial \varepsilon_j}{\partial \Lambda} = h(\Lambda) \frac{\partial \varepsilon_j}{\partial \log q_j} = \frac{\partial \log H}{\partial \Lambda} \frac{\partial \varepsilon_j}{\partial \log q_j}
$$

where we define $\log H$ as the integral of $h$:

$$
H(\Lambda) = \exp \left( \int_{\Lambda^*}^{\Lambda} h(t) dt \right)
$$

taking any fixed reference point $\Lambda^*$. We would have then $H(\Lambda^*) = 1$ by definition (it’s also important to notice that $H$ does not depend on $j$ and $q_j$).

Using this, let’s show that differential equation (69) implies:

$$
\varepsilon_j(q_j, \Lambda) = \varepsilon_j(q_j H(\Lambda), \Lambda^*)
$$

(60)
To show this result, consider the function

\[ e_j(x) = \varepsilon_j(q_j H(\Lambda))/H(x), x). \]

Taking all other variables \( \Lambda \) and \( q_j \) as fixed, only varying \( x \) between \( \Lambda^* \) and \( \Lambda \). We find that the derivative of \( e_j(x) \) w.r.t. \( x \) is zero:

\[ e_j'(x) = \frac{\partial \varepsilon_j}{\partial \Lambda} (q_j H(\Lambda)/H(x), x) - \frac{\partial \log H}{\partial \Lambda}(x) \frac{\partial \varepsilon_j}{\partial \log q_j}(q_j H(\Lambda)/H(x), x) = 0. \]

Hence \( e_j \) does not depend on \( x \). Moreover, \( e_j(\Lambda) \) corresponds to:

\[ e_j(\Lambda) = \varepsilon_j(q_j, \Lambda), \]

\[ e_j(\Lambda^*) = \varepsilon_j(q_j H(\Lambda^*), \Lambda^*) = \varepsilon_j(q_j H(\Lambda), \Lambda^*) \]

given that \( H(\Lambda^*) = 1 \) by definition of \( H \). Hence we get the equality between the last two expressions:

\[ \varepsilon_j(q_j, \Lambda) = \varepsilon_j(q_j H(\Lambda), \Lambda^*), \] which holds for any \( q_j \). Hence we have proven equation (60).

Integrating over \( q_j \) from a reference point \( q_j^* \) in the region where equality (60) holds, we obtain that demand can be written as:

\[ \frac{W_j(q_j, \Lambda)}{W_j(q_j^*, \Lambda)} = \exp \left[ \int_{q_j^*}^{q_j} \varepsilon_j(q, \Lambda) \frac{dq}{q} \right] \]

\[ = \exp \left[ \int_{q_j^*}^{q_j} \varepsilon_j(q H(\Lambda), \Lambda^*) \frac{dq}{q} \right] \]

\[ = \exp \left[ \int_{q_j^*}^{q_j H(\Lambda)} \varepsilon_j(q, \Lambda^*) \frac{dq}{q} \right] \]

\[ = \frac{W_j(q_j H(\Lambda), \Lambda^*)}{W_j(q_j^* H(\Lambda), \Lambda^*)}. \]

It shows that the effect of \( q_j \) on \( W_j \) is independent of \( \Lambda \), provided that we adjust for the shifter \( H(\Lambda) \).

Next, take a fixed reference \( q_j^* \) as given and define \( F_j \) as:

\[ F_j(\Lambda) \equiv \frac{W_j(q_j^* H(\Lambda), \Lambda^*)}{W_j(q_j^*, \Lambda)}. \]

Taking any two goods \( i \) and \( j \), we obtain:

\[ \frac{\log(F_j/F_i)}{\partial \Lambda} = h(\Lambda) \left( \varepsilon_j(q_j^* H(\Lambda), \Lambda^*) - \varepsilon_i(q_i^* H(\Lambda), \Lambda^*) \right) - \frac{\partial \log W_j}{\partial \Lambda}(q_j^*, \Lambda) + \frac{\partial \log W_i}{\partial \Lambda}(q_i^*, \Lambda) \]

\[ = h(\Lambda) \left( \varepsilon_j(q_j^* H(\Lambda), \Lambda^*) - \varepsilon_i(q_i^* H(\Lambda), \Lambda^*) \right) - \frac{\partial \log W_j}{\partial \Lambda}(q_j^*, \Lambda) + \frac{\partial \log W_i}{\partial \Lambda}(q_i^*, \Lambda) \]

\[ = 0. \]

Since \( F_j(\Lambda^*) = 1 \) for all goods \( j \), this implies that these functions \( F_j = F_i = F(\Lambda) \) is identical across all goods.
Starting with Equation (71) and combining with the properties of \( F \) above, we finally obtain:

\[
W_j(q_j, \Lambda) = \frac{W_j(q_j^*, \Lambda)}{W_j(q_j^*H(\Lambda), \Lambda^*)} \frac{W_j(q_jH(\Lambda), \Lambda^*)}{W_j(q_jH(\Lambda), \Lambda^*)} \frac{1}{F(\Lambda)} W_j(q_jH(\Lambda), \Lambda^*)
\]

Dividing by \( q_i \), this implies that normalized price must equal:

\[
\frac{p_i}{w} = \frac{1}{q_i F(\Lambda)} W_j(q_jH(\Lambda), \Lambda^*).
\]

As we assume that demand is strictly monotonic in prices, holding \( \Lambda \) constant, it can be inverted such that we can express \( q_i \) as a function of \( p_i/w \) and \( \Lambda \). Denoting \( D_i \) the inverse of \( \frac{1}{q_i} W_j(q_j, \Lambda^* \) (holding \( \Lambda^* \) fixed), we obtain:

\[
q_i = \frac{1}{H(\Lambda)} D_j(F(\Lambda)p_j/w).
\]

**Case 2** is the simplest. Suppose that \( \varepsilon_i \) is the same across \( i \)'s. Since each \( \varepsilon_i(q_i, \Lambda) \) depends only on \( q_i \) and \( \Lambda \), it must be that these elasticities only depend on \( \Lambda \), i.e.:

\[
\varepsilon_j(q_j, \Lambda) = 1 - 1/\sigma(\Lambda)
\]

for some function \( \sigma(\Lambda) \neq 1 \).

Integrating, this implies that demand can be written as:

\[
W_j(q_j, \Lambda) = A_j(\Lambda)^{-\frac{1}{\sigma(\Lambda)}} q_j^{1-\frac{1}{\sigma(\Lambda)}}
\]

for some good-specific functions \( A_j(\Lambda) \). This leads the demand function in the text:

\[
\tilde{q}_i(p_i/w, \Lambda) = A_i(\Lambda)(p_i/w)^{-\sigma(\Lambda)}.
\]

**Case 3** Suppose that \( \varepsilon_i \) is the same across \( i \)'s except for a single good \( i_0 \). Again, since each \( \varepsilon_i(q_i, \Lambda) \) depends only on \( q_i \) and \( \Lambda \) (except good \( i_0 \)), it must be that these elasticities only depend on \( \Lambda \), i.e.:

\[
\varepsilon_i(q_i, \Lambda) = \bar{\varepsilon}(\Lambda)
\]

for each good \( i \neq i_0 \), for some function \( \bar{\varepsilon}(\Lambda) \neq 0 \). In that case, Equation (58) can be rewritten:

\[
A_j + \frac{\partial \log W_{i_0}}{\partial \Lambda} \bar{\varepsilon}(\Lambda) = A_{i_0} + \frac{\partial \log W_j}{\partial \Lambda} \varepsilon_{i_0}
\]

for any good \( j \neq i_0 \). For \( i \neq j \) and \( i \neq i_0 \), we have:

\[
A_j + \frac{\partial \log W_i}{\partial \Lambda} \bar{\varepsilon}(\Lambda) = A_i + \frac{\partial \log W_j}{\partial \Lambda} \bar{\varepsilon}(\Lambda).
\]
Taking the difference, we obtain for any two goods \( j, i \neq i_0 \):

\[
\left( \frac{\partial \log W_{i_0}}{\partial \Lambda} - \frac{\partial \log W_i}{\partial \Lambda} \right) \bar{\varepsilon}(\Lambda) = (A_{i_0} - A_i) + \frac{\partial \log W_j}{\partial \Lambda} (\varepsilon_{i_0} - \bar{\varepsilon}(\Lambda)).
\]

Taking again the difference with the same expression with a fourth good \( k \) instead of \( j \), we obtain:

\[
0 = \left( \frac{\partial \log W_j}{\partial \Lambda}(q_j, \Lambda) - \frac{\partial \log W_k}{\partial \Lambda}(q_k, \Lambda) \right) (\varepsilon_{i_0}(q_{i_0}, \Lambda) - \bar{\varepsilon}(\Lambda)).
\]

Since \( \varepsilon_{i_0}(q_{i_0}, \Lambda) \neq \bar{\varepsilon}(\Lambda) \), it implies that \( \frac{\partial \log W_j}{\partial \Lambda}(q_j, \Lambda) = \frac{\partial \log W_k}{\partial \Lambda}(q_k, \Lambda) \), which must hold for any pair of good \( k \) and \( j \) except \( i_0 \). This implies that there exist some functions \( A(\Lambda) \) and \( \tilde{W}_j(q_j) \) such that \( W_j(q_j, \Lambda) = \tilde{W}_j(q_j) A(\Lambda) \) for all \( j \neq i_0 \). Since \( \frac{\partial \log W_j}{\partial \log q_j} = -\bar{\varepsilon}(\Lambda) \), we can also conclude that \( \bar{\varepsilon}(\Lambda) \) is constant and does not depend on \( \Lambda \). Thus, denoting \( \tilde{\varepsilon} = 1 - 1/\sigma \), we obtain:

\[
W_j(q_j, \Lambda) = w_j q_i^{1-\frac{1}{\sigma}} A(\Lambda),
\]

\( (63) \)

for some constant terms \( w_j \) for each \( j \neq i_0 \). We obtain the functional form in the text by inverting and expressing \( q_i \) as a function of \( \Lambda \) and \( p_i/w \).

**Combinations of cases:**

Locally, for a given \( \Lambda \) and around it, one must be in one of these three cases. A remaining question is whether demand can be a mixture of these three cases as \( \Lambda \) varies. To finish the proof of Proposition 1, we show that we cannot combine case 1 with cases 2 and 3, hence the functional form of case 1 needs to hold globally across all \( \Lambda \)'s.

**Combination of cases 1+2** Here we show that we cannot have a combination of cases 1 and 2 globally. First, note that for a given \( \Lambda \), case 1 and 2 are mutually exclusive by definition. Hence, if we have a mixture of cases 1 and 2, it must occur along different \( \Lambda \)'s. By contradiction, suppose that there exists \( \Lambda^* \) such that, at least locally,

\[
W_i(q_i, \Lambda) = W_j(q_j H(\Lambda), \Lambda^*) / F(\Lambda) \quad \text{if} \quad \Lambda < \Lambda^*
\]

\[
W_i(q_i, \Lambda) = A_i(\Lambda) - \sigma(\Lambda) q_i^{1-\frac{1}{\sigma}} \quad \text{if} \quad \Lambda > \Lambda^*
\]

By continuity, at the limit where \( \Lambda = \Lambda^* \), we must have:

\[
\frac{\partial \log W_i}{\partial \log y} = 1 - \sigma(\Lambda^*).
\]

Since it must hold for any \( i \) and any \( y \), it implies that \( \frac{\partial \log W_i}{\partial \log y} = 0 \), which contradicts our assumption that \( W_i(y_i, \Lambda) \) is not locally constant across \( y_i \) for any given \( \Lambda \).

**Combinations of cases 1+3** Here we show that we cannot have a combination of cases 1 and 3 globally, using the same arguments as above. Note again that for a given \( \Lambda \), case 1 and 3 are mutually exclusive by definition. Hence, if we have a mixture of cases 1 and 3, it must occur along different \( \Lambda \)'s.

By contradiction, suppose that there exists \( \Lambda^* \) such that, at least locally, such that for all but one good we have:

\[
W_i(q_i, \Lambda) = W_j(q_j H(\Lambda), \Lambda^*) / F(\Lambda) \quad \text{if} \quad \Lambda < \Lambda^*
\]

\[
W_i(q_i, \Lambda) = w_j H(\Lambda) q_i^{1-\frac{1}{\sigma}} \quad \text{if} \quad \Lambda > \Lambda^*.
\]
Again, by continuity, at the limit where $\Lambda = \Lambda^*$, we must have:

$$\frac{\partial \log D_i(F(\Lambda^*)y)}{\partial \log y} = 1 - \sigma.$$ 

Again, since it must hold for any $i$ and any $y$, it implies that $\frac{\partial \log W_i}{\partial \log y} = 0$, which contradicts our assumption that $W_i(y_i, \Lambda)$ is not locally constant across $y_i$ for any given $\Lambda$.

**Proposition 2**

**Preliminaries.** As for Proposition 1, it is easier to prove Proposition 2 by examining the inverse demand, i.e. normalized prices as a function of quantities $q$ (here with two aggregators $\Lambda$ and $U$).

Consider the demand system:

$$q_i = \tilde{q}_i(p_i/w, \Lambda, V)$$

where $V = V(p/w)$ is indirect utility and $\Lambda$ is an implicit function of normalized prices $p_i/w$ such that the budget constraint holds, i.e. such that:

$$\sum_i (p_i/w) \tilde{q}_i(p_i/w, \Lambda, V(p/w)) = 1.$$

If we assume that each $\tilde{q}_i(p_i/w, \Lambda, V)$ is monotonically decreasing in $\Lambda$ (here we assume a strictly negative derivative), the solution in $\Lambda$ is unique.

Since we also assume that expenditure shares $(p_i/w) \tilde{q}_i(p_i/w, \Lambda, V)$ monotonically decreases or increases with prices (holding $\Lambda$ and $V$ constant), and since we assume that for each $q_i$ there exist a vector of normalized prices such that $\tilde{q_i}(p_i/w, \Lambda(p/w), V(p/w))$, such demand can be inverted such that there exist functions $W_i$ such that:

$$q_i p_i / w = W_i(q_i, \Lambda, U)$$

i.e. such that $(p_i/w)\tilde{q}_i(p_i/w, \Lambda(p/w), V(p/w)) = W_i(\tilde{q}_i(p_i/w, \Lambda(p/w), V(p/w)), \Lambda(p/w), V(p/w))$ for any $p/w$. By definition, note also that direct and indirect utility are equal, $V(p/w) = U(q)$, when demand $q$ is evaluated at normalized prices $p/w$.

As demand $\tilde{q}_i(p_i/w, \Lambda, V)$ has a strictly negative derivative in $\Lambda$ (holding $q_i$ and $V$ constant), by the implicit theorem we can also conclude that $W_i$ has a strictly negative derivative in $\Lambda$. As in the single-aggregator case (Proposition 1), we can thus redefine $\Lambda$ as an implicit function of the vector of quantities such that the budget constraint holds, i.e. such that: $\sum_i W_i(q_i, \Lambda, U(q)) = 1$ when $V$ coincides with utility $U(q)$.

Again, in the remainder of the proof of Proposition 2, $\Lambda$ refers to a function of quantities $q$ rather than normalized prices.

**Proof of Proposition 2**

For such a demand system to be integrable (and satisfy Slutsky symmetry), there must exist a utility function $U(q)$ and another scale function such that:

$$\frac{\partial U}{\partial \log q_i} = \lambda(q) W_i(q_i, \Lambda(q), U(q)).$$

(64)
We further assume that such utility function is twice continuously differentiable. Differentiating the budget constraint $\sum_i W_i(q_i, \Lambda(q), U(q)) = 1$ implies:

$$\frac{\partial \log W_j}{\partial \log q_j} = -\sum_i \frac{\partial W_i}{\partial \Lambda} \frac{\partial \Lambda}{\partial \log q_j} - \sum_i \frac{\partial W_i}{\partial U} \frac{\partial U}{\partial \log q_j}. \quad (65)$$

Using $\frac{\partial U}{\partial \log q_i} = \lambda W_j$, we obtain:

$$S_q(q) \frac{\partial \Lambda}{\partial \log q_j} = \varepsilon_j(q_j, \Lambda) - S_U(q) \lambda(q). \quad (66)$$

where $\varepsilon_j(q_j, \Lambda) \equiv \frac{\partial \log W_i}{\partial \log q_i} \bigg|_{\Lambda}$ denotes the elasticity w.r.t. own quantity $q_j$, holding aggregators constant, where $S_q(q) \equiv -\sum_i \frac{\partial W_i}{\partial \Lambda}$ is different from zero by assumption, and where $S_U(q) \equiv -\sum_i \frac{\partial W_i}{\partial U}$.

Next, differentiating equation (64), we obtain:

$$\frac{\partial U}{\partial \log q_i \partial \log q_j} = \frac{\partial \lambda}{\partial \log q_j} W_i + \lambda \frac{\partial W_i}{\partial \Lambda} \frac{\partial \Lambda}{\partial \log q_j} + \lambda^2 \frac{\partial W_i}{\partial U} W_j. \quad (67)$$

The cross derivative is symmetric as we assume that $U$ is twice continuously differentiable. Hence, dividing by $\lambda W_i W_j$ we obtain:

$$\left( \frac{1}{W_j} \frac{\partial \log \lambda}{\partial \log q_j} \right) + \frac{\partial \log W_i}{\partial \Lambda} \left( \frac{1}{W_j} \frac{\partial \Lambda}{\partial \log q_j} \right) + \lambda \frac{\partial \log W_i}{\partial U} =$$

$$\left( \frac{1}{W_i} \frac{\partial \log \lambda}{\partial \log q_i} \right) + \frac{\partial \log W_j}{\partial \Lambda} \left( \frac{1}{W_i} \frac{\partial \Lambda}{\partial \log q_i} \right) + \lambda \frac{\partial \log W_j}{\partial U}.$$  

Incorporating the expression from (66), this is equivalent to:

$$\left( \frac{S_q}{W_j} \frac{\partial \log \lambda}{\partial \log q_j} \right) + \frac{\partial \log W_i}{\partial \Lambda} (\varepsilon_j - S_U \lambda) + \lambda \frac{\partial \log W_i}{\partial U} =$$

$$\left( \frac{S_q}{W_i} \frac{\partial \log \lambda}{\partial \log q_i} \right) + \frac{\partial \log W_j}{\partial \Lambda} (\varepsilon_i - S_U \lambda) + \lambda \frac{\partial \log W_j}{\partial U}.$$  

Define $A_i(q) = \frac{S_q}{W_i} \frac{\partial \log \lambda}{\partial \log q_i} + \frac{\partial \log W_i}{\partial \Lambda} - \lambda \frac{\partial \log W_i}{\partial U}$ we get an expression that is very similar to the single-aggregator case:

$$A_j(q) + \frac{\partial \log W_i}{\partial \Lambda}(q_j, \Lambda, U) \varepsilon_j(q_j, \Lambda, U) = A_i(q) + \frac{\partial \log W_j}{\partial \Lambda}(q_j, \Lambda, U) \varepsilon_i(q_i, \Lambda, U) \quad (68)$$

and holds for any $i \neq j$.

Unlike the previous Proposition, here we directly assume that $\varepsilon_i(q_i, \Lambda, U)$ takes at least two different values across goods $i$, almost everywhere, even if we exclude any one good.

In this case, even if we exclude a single good $j$, there exists a vector $x_i(q)$ such that $\sum_i x_i = 0$ and $\sum_i x_i x_i \neq 0$. Multiplying Equation (68) by $x_i(q)$ and summing up across goods $i$ (for a given $j$), we obtain:

$$\left( \sum_i x_i \frac{\partial \log W_i}{\partial \Lambda} \right) \varepsilon_j = \left( \sum_i x_i A_i \right) + \left( \sum_i x_i \varepsilon_i \right) \frac{\partial \log W_j}{\partial \Lambda}. \quad \text{50}$$
As \( \sum \varepsilon_i x_i \neq 0 \), we obtain that there exists two functions \( h(q) \) and \( m(q) \) such that:

\[
\frac{\partial \log W_j}{\partial \Lambda}(q_j, \Lambda, U) = h(q) \varepsilon_j(q_j, \Lambda, U) + m(q).
\]

In particular, this holds also for any pair of goods \( i \) and \( j \). Taking the difference, we get:

\[
\frac{\partial \log W_j}{\partial \Lambda}(q_j, \Lambda, U) - \frac{\partial \log W_i}{\partial \Lambda}(q_i, \Lambda, U) = h(q) \left( \varepsilon_j(q_j, \Lambda, U) - \varepsilon_i(q_i, \Lambda, U) \right)
\]

Take two goods for which \( \varepsilon_i \neq \varepsilon_j \). Note that the left-hand side only depends on \( q_j, q_i \) and \( \Lambda \). This implies that \( h(q) \) can be written as a function of \( q_j, q_i, \Lambda \) and \( U \) only.

We can also find a third good \( i' \) such that \( \varepsilon_{i'} \neq \varepsilon_i \) and \( \varepsilon_{i'} \neq \varepsilon_j \). Applying the same argument, it must be that \( h \) can be written as just a function of \( \Lambda \) and \( U \), so we now denote \( h \) as: \( h = h(\Lambda, U) \).

Taking again a derivative in \( \log q_j \), holding \( \Lambda \) and \( U \) constant, and noticing that the cross derivative is symmetric, \( \frac{\partial \varepsilon_j}{\partial \Lambda} = \frac{\partial \log W_j}{\partial \log q_j \partial \Lambda} = \frac{\partial \log W_i}{\partial \Lambda \partial \log q_j} \), we obtain:

\[
\frac{\partial \varepsilon_j}{\partial \Lambda} = h(\Lambda, U) - \frac{\partial \log H}{\partial \log q_j} \quad \frac{\partial \log H}{\partial \Lambda} \quad \frac{\partial \varepsilon_j}{\partial \log q_j}
\]

where we define \( \log H \) as the integral of \( h \), for a given \( U \):

\[
H(\Lambda, U) = \exp \left( \int_{\Lambda^*}^{\Lambda} h(t, U) dt \right)
\]

taking any fixed reference point \( \Lambda^* \). We would have then \( H(\Lambda^*, U) = 1 \) by definition (it’s also important to notice that \( H \) does not depend on \( j \) and \( q_j \)).

Using this, let’s show that differential equation (69) implies:

\[
\varepsilon_j(q_j, \Lambda, U) = \varepsilon_j(q_j H(\Lambda, U), \Lambda^*, U)
\]

To show this result, consider the function

\[
e_j(x) = \varepsilon_j(q_j H(\Lambda, U)/H(x, U), x, U)
\]

Taking all other variables \( \Lambda \), \( U \) and \( q_j \) as fixed, only varying \( x \) between \( \Lambda^* \) and \( \Lambda \). We find that the derivative of \( e_j(x) \) w.r.t. \( x \) is zero:

\[
e'_j(x) = \frac{\partial \varepsilon_j}{\partial \Lambda} \left( q_j H(\Lambda, U)/H(x, U), x, U \right) - \frac{\partial \log H}{\partial \Lambda}(x, U) \frac{\partial \varepsilon_j}{\partial \log q_j} \left( q_j H(\Lambda, U)/H(x, U), x, U \right) = 0
\]

Hence \( e_j \) does not depend on \( x \). Moreover, \( e_j(\Lambda) \) corresponds to: \( e_j(\Lambda) = \varepsilon_j(q_j, \Lambda, U) \), while \( e_j(\Lambda^*) \) is such that:

\[
e_j(\Lambda^*) = \varepsilon_j(q_j H(\Lambda, U)/H(\Lambda^*, U), \Lambda^*, U) = \varepsilon_j(q_j H(\Lambda, U), \Lambda^*, U)
\]

given that \( H(\Lambda^*, U) = 1 \) by definition of \( H \). Hence we get the equality between the last two expressions: \( \varepsilon_j(q_j, \Lambda, U) = \varepsilon_j(q_j H(\Lambda, U), \Lambda^*, U) \), which holds for any \( q_j \). Thus we have proven equation (70).

Integrating over \( q_j \) from a reference point \( q_j^* \) in the region where equality (70) holds, we obtain
that demand can be written as:

$$\frac{W_j(q_j, \Lambda, U)}{W_j(q_j^*, \Lambda, U)} = \exp \left[ \int_{q_j}^{q_j^*} \varepsilon_j(q, \Lambda, U) \frac{dq}{q} \right] = \exp \left[ \int_{q_j^*}^{q_j} \varepsilon_j(q, \Lambda, U) \frac{dq}{q} \right] = \frac{W_j(q_j, \Lambda, U)}{W_j(q_j^* H(\Lambda, U), \Lambda^*, U)}.$$ 

It shows that the effect of $q_j$ on $W_j$ is independent of $\Lambda$ and $U$, provided that we adjust for the shifter $H(\Lambda, U)$.

Next, take a fixed reference $q_j^*$ as given and define $F_j$ as:

$$F_j(\Lambda, U) \equiv \frac{W_j(q_j^* H(\Lambda, U), \Lambda^*, U)}{W_j(q_j^*, \Lambda, U)}.$$ 

Taking any two goods $i$ and $j$, we obtain:

$$\log \left( \frac{F_j/\Lambda_i}{\partial \Lambda} \right) = h(\Lambda, U) \left( \varepsilon_j(q_j^* H(\Lambda, U), \Lambda^*, U') - \varepsilon_i(q_i^* H(\Lambda, U), \Lambda^*, U) \right) - \frac{\partial \log W_j}{\partial q_j}(q_j^*, \Lambda, U) + \frac{\partial \log W_i}{\partial q_j}(q_i^*, \Lambda, U)$$

$$= h(\Lambda, U) \left( \varepsilon_j(q_j^*, \Lambda, U) - \varepsilon_i(q_i^*, \Lambda, U) \right) - \frac{\partial \log W_j}{\partial q_j}(q_j^*, \Lambda, U) + \frac{\partial \log W_i}{\partial q_j}(q_i^*, \Lambda, U)$$

$$= 0.$$ 

Since $F_j(\Lambda^*, U) = 1$ for all goods $j$, this implies that these functions $F_j = F_i = F(\Lambda, U)$ are identical across all goods.

Starting with Equation (71) and combining with the properties of $F$ above, we finally obtain:

$$W_j(q_j, \Lambda, U) = \frac{W_j(q_j^* H(\Lambda, U), \Lambda^*, U)}{W_j(q_j^* H(\Lambda, U), \Lambda^*, U)} \times W_j(q_j H(\Lambda, U), \Lambda^*, U)$$

$$= \frac{1}{F(\Lambda, U)} W_j(q_j H(\Lambda, U), \Lambda^*, U).$$ 

Dividing by $q_i$, this implies that normalized price must equal:

$$\frac{p_i}{w} = \frac{1}{q_i F(\Lambda, U)} W_j(q_j H(\Lambda, U), \Lambda^*, U).$$ 

As we assume that demand is strictly monotonic in prices, holding $\Lambda$ and $U$ constant, it can be inverted such that we can express $q_i$ as a function of $p_i/w$ and $\Lambda$. Denoting $D_i$ the inverse of $\frac{1}{q_i} W_j(q_j, \Lambda^*, U)$
w.r.t. \( q_i \) (holding \( U \) constant and holding \( \Lambda^* \) fixed), we obtain the expression in Proposition 2:

\[
q_i = \frac{1}{H(\Lambda, V)} D_j(F(\Lambda, V)p_j/w, V). \tag{71}
\]

**Proof of Proposition 3**

Define \( \widetilde{U}(q, \Lambda) \) as:

\[
\widetilde{U}(q, \Lambda) = \sum_i u_i(H(\Lambda)q_i) - \int_{\Lambda_0}^{\Lambda} F(\Lambda) H'(\Lambda) d\Lambda
\]

where:

\[
u_i(q_i) = \int_{q_i=0}^{q_i} D_i^{-1}(x) dx
\]

and \( u'_i = D_i^{-1} \). Next, \( \Lambda \) can be defined as an implicit function of \( q \) such that:

\[
\sum_i q_i u'_i(H(\Lambda)q_i)/F(\Lambda) = 1 \tag{72}
\]

As in Propositions 1 and 2, the remainder of the proof refers to \( \Lambda \) as a function of \( q \) rather than normalized prices \( p/w \).

We proceed in three steps. First we show that equation (72) admits a solution \( \Lambda(q) \) for each \( q \) and that this solution is unique. Second we show that utility defined as \( U(q) = \widetilde{U}(q, \Lambda(q)) \) is well-behaved and quasi-concave. Finally, we show that maximizing \( U \) leads to the demand function in the text, and that the single aggregator \( \Lambda \) is also well defined and coincides with \( \Lambda \) for optimal consumption baskets.

**Step 1: Implicit function \( \Lambda(q) \).** Here we show that for any vector \( q \) of consumption, there is a unique \( \Lambda \) such that equation (72) holds.

First, using part ii) of restrictions [A3], we can see that the elasticity of \( D_i(F(\Lambda)y_i)/H(\Lambda) \) w.r.t. \( \Lambda \) is given by \( \varepsilon_F \varepsilon_{D_i} - \epsilon_H \) which is assumed to be negative, hence it strictly decreases with \( \Lambda \). Symmetrically, we obtain that \( u_i'(H(\Lambda)q_i)/F(\Lambda) \) also strictly decreases with \( \Lambda \): its elasticity w.r.t. \( \Lambda \) is \( \varepsilon_H/\varepsilon_{D_i} - \varepsilon_F \), which is also negative given that \( \varepsilon_{D_i} \) is negative and \( \varepsilon_H - \varepsilon_{D_i}\varepsilon_F \) is positive. Adding up across goods, we obtain that the left-hand side of equation (72) decreases strictly with \( \Lambda \). This implies that the solution to equation (72) is unique (if it exists).

Existence is then guaranteed using condition [A3]-iii), which we can symmetrically reformulate in terms of quantities. We assume that, for any good \( i \) and \( y_i > 0 \), there exists \( \Lambda \in \mathbb{R} \) such that: \( y_i D_i(y_iF(\Lambda))/H(\Lambda) = 1/N \). Using \( u'_i = D_i^{-1} \), note that this equality is equivalent to \( 1/N = 1/(Ny_i)u'_i(1/(Ny_i)H(\Lambda))/F(\Lambda) \). Hence, denoting \( q_i = 1/Ny_i \), we obtain that for any good \( i \) and \( q_i > 0 \), there exists \( \Lambda \in \mathbb{R} \) such that:

\[
q_i u'_i(q_iH(\Lambda))/F(\Lambda) = 1/N.
\]

For a given vector of quantities \( q \), for each good we obtain a \( \Lambda \) such that the equality above holds. Taking the maximum of the \( \Lambda \)'s obtained across the \( N \) goods (and using the monotonicity property

---

\[39\]Recall that \( D_i \) is strictly decreasing unless \( D_i = 0 \). As noted in the text, as an abuse of notation, we define \( D_i^{-1}(0) = a_i \) if \( D_i(y) = 0 \) for all \( y \geq a_i \) (which yields a choke price) and \( D_i^{-1}(x) = 0 \) for all \( x \geq b_i \) if \( D_i(0) = b_i \).
in $\Lambda$ described just above), we obtain a $\Lambda_{\text{max}}$ such that

$$\sum_i q_i u'_i (q_i H(\Lambda_{\text{max}}))/F(\Lambda_{\text{max}}) \leq 1.$$ 

For the same vector $q$, by taking the minimum of such $\Lambda$’s across goods, we obtain a $\Lambda_{\text{min}}$ such that

$$\sum_i q_i u'_i (q_i H(\Lambda_{\text{min}}))/F(\Lambda_{\text{min}}) \geq 1.$$ 

As the left hand side of this expression is continuous in $\Lambda$, the intermediate-value theorem ensures that a solution to equation (72) exists between $\Lambda_{\text{min}}$ and $\Lambda_{\text{max}}$.

Finally, note that the derivative of the left-hand side of equation (72) is strictly negative. Using the implicit function theorem, we can thus obtain the derivatives of $\Lambda$ w.r.t. $q$ as described below.

**Step 2: Quasi-concavity.** The second step is to show that utility defined as $U(q) = \tilde{U}(q, \Lambda(q))$ is quasi-concave. First, we need to compute the first and second derivatives.

**Derivatives of the aggregator $\Lambda$.** Here we consider the properties of $\Lambda(q)$, the solution of equation (72). Taking the derivative of equation (72), we get:

$$\sum_i q_i u'_i (H(\Lambda)q_i)/F(\Lambda) = 1$$

$$\frac{\partial \Lambda}{\partial q_i} \left[ H' \sum_j q^2_j u''_j - F' \right] + \left[ u'_i + Hq_i u''_i \right] = 0$$

and thus:

$$\frac{\partial \Lambda}{\partial q_i} = \frac{u'_i + Hq_i u''_i}{\Delta(q)}$$

with $\Delta(q) \equiv F' - H' \sum_i q^2_i u''_i$.

We can verify that $\Delta(q)$ is positive. Note that $\frac{u'_i}{u''_i Hq_i} = \varepsilon_{D_i}$, the elasticity of function $D_i$. Thus, we obtain:

$$\Delta(q) = F' - H' \sum_i q^2_i u''_i$$

$$= (F/\Lambda) \left( \varepsilon_F - \varepsilon_H \sum_i q^2_i u''_i / \sum_i q_i u'_i \right)$$

$$= (F/\Lambda) \left( \varepsilon_F - \varepsilon_H \sum_i q_i u'_i (1/\varepsilon_{D_i}) / \sum_i q_i u'_i \right).$$

Recall that $u'_i > 0$ and that assumption [A3]-ii) imposes: $\varepsilon_F \varepsilon_{D_i} < \varepsilon_H$ for all $i$. Since we also assume downward slopping demand, $\varepsilon_{D_i} < 0$, this implies $\varepsilon_F > \varepsilon_H/\varepsilon_{D_i}$ for all $i$ and therefore $\Delta > 0$. This implies that the derivatives of $\Lambda$ are always well defined. Also, knowing that $\Delta$ is positive will be useful again below.
Derivatives of utility $U$. The first derivatives are:

$$\frac{\partial U}{\partial q_i} = H u'_i (H q_i) + \frac{\partial \Lambda}{\partial q_i} \left[ H' \sum_i q_i u'_i (H q_i) - H' F \right] = H u'_i (H q_i)$$

where the term in brackets is null for any $q$, thanks to condition (72). Second derivatives are then:

$$\frac{\partial^2 U}{\partial q_i^2} = \frac{\partial \Lambda}{\partial q_i} (u'_i + H q_i u''_i) H' + H^2 u''_i$$

$$\frac{\partial^2 U}{\partial q_i \partial q_j} = \frac{\partial \Lambda}{\partial q_j} (u'_i + H q_i u''_i) H'$$

and thus, incorporating the derivatives in $\Lambda$, we obtain:

$$\frac{\partial^2 U}{\partial q_i^2} = (u'_i + H q_i u''_i)^2 H' / \Delta + H^2 u''_i$$

$$\frac{\partial^2 U}{\partial q_i \partial q_j} = (u'_i + H q_i u''_i) (u'_j + H q_j u''_j) H' / \Delta.$$

**Negative semi-definiteness.** To show that utility is quasi-concave, we need to show that the bordered Hessian is semi-definite negative, i.e we need to show:

$$\sum_{i,j} t_i t_j \frac{\partial^2 U}{\partial q_i \partial q_j} = \left( \sum_i t_i (u'_i + H q_i u''_i) \right)^2 H' / \Delta + \sum_i t_i^2 H^2 u''_i < 0$$

for any vector $t \in \mathbb{R}^N$ such that:

$$\sum_i t_i \frac{\partial U}{\partial q_i} = \sum_i t_i H u'_i = 0.$$

The objective function above is homogeneous of degree 2. We can thus normalize the sum $\sum_i t_i (u'_i + H q_i u''_i)$ up to any constant without loss of generality.

The first step is to find the optimal vector of $t_i$’s that maximizes the left-hand side of the inequality above. It is equivalent to consider the maximization:

$$\max \left\{ \sum_i t_i^2 u''_i \right\}$$

under the constraint: $\sum_i t_i (u'_i + H q_i u''_i) = \text{constant}$ and $\sum_i t_i u'_i = 0$. The first-order condition is: $2 u''_i t_i = \mu_1 u'_i + \mu_2 (u'_i + H q_i u''_i)$ where $\mu_1$ and $\mu_2$ are the Lagrange multipliers for the two constraints. This leads to $t_i$ being proportional to:

$$t_i \sim \frac{u'_i}{H u''_i} + \mu q_i$$

for some $\mu$ (note that the second-order conditions are satisfied as the objective function is concave: $u''_i < 0$ for all goods $i$). Given that we must have $0 = \sum_i t_i u'_i = \sum_i \frac{u_i'^2}{H u''_i} + \mu \sum_i q_i u'_i$, $\mu$ must correspond
The term in parentheses is the same on the left and on the right. This term is negative iff:

\[ \Delta > 0, \]

where \( \varepsilon_{Di} = \frac{u_i'}{Hu_i'} \) and \( \bar{\varepsilon}_D \) is its weighted average (weighted by \( q_iu_i' \)).

Next, using the optimal \( t_i = \frac{u_i'}{Hu_i'} - \bar{\varepsilon}_Dq_i = q_i\bar{\varepsilon}_{Di} - q_i\bar{\varepsilon}_D \), a sufficient and necessary condition for negative semi-definiteness is:

\[
\left( \sum_i (q_i\varepsilon_{Di} - q_i\bar{\varepsilon}_D)(u_i' + q_iHu_i'') \right)^2 \frac{H'}{\Delta} + H^2 \sum_i (q_i\varepsilon_{Di} - q_i\bar{\varepsilon}_D)^2 u_i'' < 0.
\]

Since \( \Delta > 0 \), this condition can be rewritten:

\[
\left( \sum_i (q_i\varepsilon_{Di} - q_i\bar{\varepsilon}_D)(u_i' + q_iHu_i'') \right)^2 H' < -H^2 \Delta \left( \sum_i (q_i\varepsilon_{Di} - q_i\bar{\varepsilon}_D)^2 u_i'' \right).
\]

The term in parentheses is the same on the left and on the right. This term is negative iff:

\[
\sum_i q_iu_i' - \bar{\varepsilon}_D H \sum_i q_i^2 u_i'' < 0 \iff \sum_i q_iu_i' < \left( \frac{\sum_i q_iu_i'\varepsilon_{Di}'}{\sum_i q_iu_i'} \right) \left( \sum_i q_iu_i'/\varepsilon_{Di} \right).
\]
This last inequality is satisfied as long as price elasticity are not equal across all goods: the left hand side corresponds to a harmonic average while the right-hand-side corresponds to an arithmetic average of a positive variable $-\varepsilon_{Di} > 0$.

Hence, using $\sum_i q_i u_i' - \bar{\varepsilon}_D H \sum_i q_i^2 u_i'' < 0$ and also that $\Delta \equiv F' - H' \sum_i q_i^2 u_i''$ the previous inequality is equivalent to:

\[
\Leftrightarrow H' \left( \sum_i q_i u_i' - \bar{\varepsilon}_D H \sum_i q_i^2 u_i'' \right) > \bar{\varepsilon}_D H \Delta
\]

\[
\Leftrightarrow H' \left( \sum_i q_i u_i' - \bar{\varepsilon}_D H \sum_i q_i^2 u_i'' \right) > \bar{\varepsilon}_D H \left( F' - H' \sum_i q_i^2 u_i'' \right)
\]

\[
\Leftrightarrow H' \sum_i q_i u_i' > \bar{\varepsilon}_D H F'.
\]

Given that $F = \sum_i q_i u_i'$, this inequality is equivalent to:

\[
\Leftrightarrow H' F > \bar{\varepsilon}_D H F'
\]

\[
\Leftrightarrow \varepsilon H > \bar{\varepsilon}_D \varepsilon_F.
\]

This holds, given that $\bar{\varepsilon}_D$ is a weighted average of $\varepsilon_{Di}$, and $\varepsilon_{Di} \varepsilon_F < \varepsilon_H$ is assumed in part ii) of restrictions [A3] for each good $i$.

**Step 3: Marshallian demand and price aggregator.** Maximizing $U(q)$ under the budget constraint $\sum_i p_i q_i = w$ leads to:

\[
\frac{\partial U}{\partial q_i} = H(\Lambda) u_i'(H(\Lambda) q_i) = \mu p_i
\]

where $\mu$ henceforth denotes the Lagrange multiplier associated with the budget constraint. Summing across goods, we can see that $\mu$ is such that:

\[
\mu = \frac{1}{w} \sum_i \mu p_i q_i = \frac{1}{w} \sum_i H q_i u_i'(H q_i) = \frac{H(\Lambda) F(\Lambda)}{w}.
\]

Using $H(\Lambda) u_i'(H(\Lambda) q_i) = \mu p_i$, we obtain:

\[
u_i'(H(\Lambda) q_i) = \frac{\mu p_i}{H(\Lambda)} = \frac{F(\Lambda) p_i}{w}
\]

and thus, given the definition of $u_i'$:

\[
H(\Lambda) q_i = D_i(\mu p_i / H(\Lambda)) = D_i(F(\Lambda) p_i / w)
\]

and:

\[
q_i = D_i(F(\Lambda) p_i / w) / H(\Lambda)
\]
The final step is to show that \( \Lambda \) can be implicitly defined as a function of all normalized prices \( p_i/w \). To see this, notice that \( q_i \) must satisfy the budget constraint:

\[
w = \sum_i q_i p_i = \sum_i p_i D_i(F(\Lambda) p_i/w) / H(\Lambda)
\]

which can be rewritten:

\[
\sum_i (p_i/w) D_i(F(\Lambda) p_i/w) / H(\Lambda) = 1.
\]

The solution of this equation in \( \Lambda \) is unique, which shows that we can alternatively define \( \Lambda \) as a function of normalized prices \( p_i/w \). To prove that there is a unique solution, we can follow the same approach and assumptions as in Step 1 above: condition \([A3]-\text{ii})\) ensures uniqueness while condition \([A3]-\text{iii})\) provides existence.

**Alternative proof of Proposition 3 using the Slutsky Matrix** Alternatively, it is possible to prove Proposition 3 by showing that the Slutsky matrix is symmetric and negative semi-definite, and then apply Hurwicz and Uzawa (1971) theorem. This is the approach taken by Matsuyama and Ushchev (2017) for the homothetic case. A similar approach can be extended here to the non-homothetic case (see a previous working paper version, Fally 2018).

**From direct to indirect utility** We start from the following geometric equality that applies to any strictly monotonic mapping \( T \):

\[
\int_{q_0}^{q_1} T^{-1}(q) dq + T^{-1}(q_0) q_0 = - \int_{y_0}^{y_1} T(y) dy + T(y_1) y_1
\]

with \( q_0 = T(y_0) \) and \( q_1 = T(y_1) \). Applying this formula to \( T = D_i, q_1 = H(\Lambda) q_i \) and \( y_1 = F(\Lambda) p_i/w \), we obtain:

\[
\int_{q_{0i}}^{H(\Lambda) q_i} D_i^{-1}(q) dq = - \int_{y_{0i}}^{F(\Lambda) p_i/w} D_i(y) dy + D_i(F(\Lambda) p_i/w) F(\Lambda) p_i/w - y_0 q_{0i}
\]

with \( y_{0i} = D_i(q_{0i}) \) for each \( i \). Moreover, note that we have:

\[
\sum_i (p_i/w) D_i(F(\Lambda) p_i/w) = H(\Lambda).
\]

Applying these equalities to the expression for direct utility provided in the text, we obtain (indirect) utility as a function of normalized prices:

\[
U = \sum_i u_i(H(\Lambda) q_i) - \int_{\Lambda_0}^{\Lambda} F(\Lambda) H'(\Lambda) d\Lambda
\]

\[
= \sum_i \int_{q=0}^{H(\Lambda) q_i} D_i^{-1}(x) dx - \int_{\Lambda_0}^{\Lambda} F(\Lambda) H'(\Lambda) d\Lambda
\]

\[
= - \sum_i \int_{y_{0i}}^{F(\Lambda) p_i/w} D_i(y) dy + \sum_i D_i(F(\Lambda) p_i/w) F(\Lambda) p_i/w - \int_{\Lambda_0}^{\Lambda} F(\Lambda) H'(\Lambda) d\Lambda - \sum_i y_{0i} q_{0i}
\]
\[
\begin{align*}
= & -\sum_i \int_{y_{0i}}^F(\Lambda)p_i/w \, D_i(y) \, dy + F(\Lambda)H(\Lambda) - \int_{\Lambda_0}^\Lambda F(\Lambda)H'(\Lambda) \, d\Lambda - \sum_i y_{0i}q_{0i} \\
= & -\sum_i \int_{y_{0i}}^F(\Lambda)p_i/w \, D_i(y) \, dy + \int_{\Lambda_0}^\Lambda F(\Lambda)'H(\Lambda) \, d\Lambda + F(\Lambda_0)H(\Lambda_0) - \sum_i y_{0i}q_{0i} \\
= & -\sum_i \int_{y_{0i}}^F(\Lambda)p_i/w \, D_i(y) \, dy + \int_{\Lambda_0}^\Lambda F(\Lambda)'H(\Lambda) \, d\Lambda + g_0
\end{align*}
\]

where \( g_0 = F(\Lambda_0)H(\Lambda_0) - \sum_i y_{0i}q_{0i} \) is a constant term.

A counter-example when condition [A3]-ii) fails.

Here I show that we can find a case where conditions ii) fails and where the Slutsky substitution matrix is not semi-definite negative, thus proving that condition ii) cannot be entirely waived.

Suppose that \( F(\Lambda) = \Lambda \) (no problem arises when \( F \) is locally constant) and that we have two goods 1 and 2, where \( \varepsilon_{D1} < \varepsilon_H \) while \( \varepsilon_{D2} > \varepsilon_H \) for the other good, i.e. \( \varepsilon_H \in (\varepsilon_{D1}, \varepsilon_{D2}) \). In particular, to fix ideas, supposed that all elasticities are constant, with \( \varepsilon_H = \frac{\varepsilon_{D2} + \varepsilon_{D1}}{2} \equiv -\kappa < 0 \) and denote \( \delta \equiv \varepsilon_{D2} - \varepsilon_H = \varepsilon_H - \varepsilon_{D1} > 0 \). Denote by the expenditure share of product 1 as \( \frac{\bar{\varepsilon}_1}{\bar{q}_1} \) and the expenditure share of good 2 as \( \frac{\bar{\varepsilon}_2}{\bar{q}_2} \) such that \( \bar{\varepsilon}_D - \bar{\varepsilon}_H = \varepsilon \delta \). While elasticities are constant, we can still adjust the demand shifter for each good to obtain the desired market shares (hence \( \varepsilon \) can be chosen independently from the elasticities).

The off-diagonal coefficients of the Slutsky substitution matrix are then:

\[
\begin{align*}
\frac{s_{12}p_1p_2}{w} = & -a_1a_2(\varepsilon_{D1} - \varepsilon_H)(\varepsilon_{D2} - \varepsilon_H) \frac{\varepsilon_D - \varepsilon_H}{\varepsilon_D - \varepsilon_H} + a_1a_2\varepsilon_H = -\frac{(1 - \varepsilon^2)\delta^2}{4\varepsilon \delta} - \frac{(1 - \varepsilon^2)\kappa}{4} \\
\frac{s_{11}p_1^2}{w} = & a_1\varepsilon_{D1} - \frac{a_1^2(\varepsilon_{D1} - \varepsilon_H)^2}{\varepsilon_D - \varepsilon_H} + a_1^2\varepsilon_H = -\frac{(1 - \varepsilon)(\kappa + \delta)}{2} + \frac{(1 - \varepsilon^2)\delta^2}{4\varepsilon \delta} - \frac{(1 - \varepsilon^2)\kappa}{4} \\
\frac{s_{22}p_2^2}{w} = & a_2\varepsilon_{D2} - \frac{a_2^2(\varepsilon_{D2} - \varepsilon_H)^2}{\varepsilon_D - \varepsilon_H} + a_2^2\varepsilon_H = -\frac{(1 + \varepsilon)(\kappa - \delta)}{2} + \frac{(1 + \varepsilon^2)\delta^2}{4\varepsilon \delta} - \frac{(1 + \varepsilon^2)\kappa}{4}.
\end{align*}
\]

One can see that the substitution coefficients become very large as \( \varepsilon \) approach zero (because some of the terms have \( \varepsilon \) in the denominator). Moreover, if we denote by \( \Sigma \) the matrix with coefficients \( s_{ij}p_ip_j/w \), we obtain:

\[
\lim_{\varepsilon \to 0^+} 4\varepsilon \Sigma = \begin{pmatrix}
+\delta & -\delta \\
-\delta & +\delta
\end{pmatrix}.
\]

This matrix is semi-definite positive: \( x^T4\varepsilon \Sigma x = \delta^2(x_1 - x_2)^2 \geq 0 \). By continuity, when \( \varepsilon \) is small enough, the substitution matrix with coefficient \( s_{ij} \) is semi-definite positive, which is not consistent with a rational demand system.
Proof of Proposition 4

Suppose that demand can be written:

\[ q_i = G_i(\Lambda)^{1-\sigma(\Lambda)} (p_i/w)^{-\sigma(\Lambda)} \]

with \( \Lambda \) implicitly defined by \( \sum_i [G_i(\Lambda)p_i/w]^{1-\sigma(\Lambda)} = 1 \).

The goal is to show that these equations:

\[
\left[ \sum_i (G_i(\Lambda)p_i/w)^{1-\sigma(\Lambda)} \right]^{\frac{1}{1-\sigma(\Lambda)}} = 1 \quad (73)
\]

\[
\left[ \sum_i (G_i(U)/q_i)^{\frac{1-\sigma(U)}{\sigma(U)}} \right]^{\frac{\sigma(U)}{1-\sigma(U)}} = 1 \quad (74)
\]

have a unique solution in \( \Lambda \) and \( U \) respectively. To do so, we show that the left-hand side of each of these equations strictly increase in \( \Lambda \) and \( U \) around the solution, showing that the left-hand side can be equal to unity only once.

We distinguish two cases, depending on whether elasticity \( \sigma(\Lambda) \) increases with \( \Lambda \). If the first case we assume that \( G_i(\Lambda) \) strictly increases with \( \Lambda \). In the second case, we impose condition ii).

1) In the first case, suppose that \( \sigma(\Lambda) \) increases with \( \Lambda \) and that \( G_i(\Lambda) \) strictly increases with \( \Lambda \). The equation above in \( \Lambda \) is equivalent to:

\[ \sum_i (G_i(\Lambda)p_i/w)^{1-\sigma(\Lambda)} = 1. \]

If \( \sigma(\Lambda) \in (0, 1) \), each term \( G_i(\Lambda)p_i/w \) in the summation increases in \( \Lambda \) and has to be smaller than unity. Hence, if \( 1 - \sigma(\Lambda) \) decreases with \( \Lambda \), the left-hand side of this expression is strictly increasing with \( \Lambda \). The same holds if we raise the whole expression on the left-hand side to the power \( \frac{1}{1-\sigma(\Lambda)} \).

If \( \sigma(\Lambda) > 1 \), each term \( G_i(\Lambda)p_i/w \) in the summation increases in \( \Lambda \) and has to be larger than unity. Hence, if \( 1 - \sigma(\Lambda) \) decreases with \( \Lambda \) (i.e. becomes more positive), the left-hand side of this expression is strictly decreasing in \( \Lambda \). The inverse holds if we raise the whole expression on the left-hand side to the power \( \frac{1}{1-\sigma(\Lambda)} < 0 \).

Now consider the equation:

\[ \sum_i \left( G_i(U)/q_i \right)^{\frac{1-\sigma(U)}{\sigma(U)}} = 1. \]

If \( \sigma(\Lambda) \in (0, 1) \), the exponent \( \frac{1-\sigma(U)}{\sigma(U)} \) is positive and decreases with \( U \). The term within parenthesis increases in \( U \). Moreover, each summation term has to be smaller than unity. Hence, as \( U \) increases, each summation term increases (strictly) with \( U \). The same holds if we raise the whole expression on the left-hand side to the power \( \frac{\sigma(U)}{1-\sigma(U)} \).

If \( \sigma(\Lambda) > 1 \), the exponent \( \frac{1-\sigma(U)}{\sigma(U)} \) is negative and decreases with \( U \). The term within parenthesis increases in \( U \). Moreover, each summation term has to be larger than unity. Hence, as \( U \) increases, each summation term decreases (strictly) with \( U \). If we raise the whole expression on the left-hand side to the power \( \frac{\sigma(U)}{1-\sigma(U)} \), we obtain a strictly increasing function of \( U \).
2) In the second case, we assume that $\sigma(\Lambda)$ decreases with $\Lambda$ and that, around each solution $\Lambda_0$ of equation (73), there exists a set of $\alpha_i$ such that $\sum \alpha_i = 1$ and such that $G_i(\Lambda) \alpha_i^{\frac{1}{1-\sigma(\Lambda)}}$ increases in $\Lambda$.

Define $K_i(\Lambda) = G_i(\Lambda) \alpha_i^{\frac{1}{1-\sigma(\Lambda)}}$ The left-hand side of equation (73) can then be rewritten:

$$\left[ \sum \alpha_i(K_i(\Lambda) p_i/w)^{1-\sigma(\Lambda)} \right]^{\frac{1}{1-\sigma(\Lambda)}}$$

To show that it strictly increases in $\Lambda$, we use Lemma 9 discussed in the next appendix section. We obtain that the left-hand side of the above equation decreases with $\sigma$, which itself decreases with $\Lambda$. Moreover, the term $K_i(\Lambda)$ strictly increases in $\Lambda$, by assumption, hence the whole left term strictly increases with $\Lambda$.

We can again use the same approach to show that the left-hand side of (78) increases strictly with $U$. This is equivalent to showing that the following expression strictly increases in $U$:

$$\left[ \sum \alpha_i\left( K_i(U)/q_i \right)^{\frac{\sigma(U)-1}{\sigma(U)}} \right]^{\frac{\sigma(U)}{\sigma(U)-1}}$$

Each exponent $\frac{1-\sigma(U)}{\sigma(U)}$ increases in $U$ and each term $K_i(U)$ strictly increases with $U$. With Lemma 9 again, we obtain that the whole term strictly increases with $U$.

Hence, in both cases, $\Lambda$ and $U$ are well defined by equations (73) and (78) which admit no more than one solution. This implicitly defines utility $U$ as a function of $q_i$. It is straightforward to see that such utility function is quasi-concave in $q$: indifference curves have the same shape as CES indifference curves, holding $\sigma = \sigma(U)$ constant.

Consumption quantities $q$ chosen to maximize $U$ would satisfy the following first-order conditions:

$$\frac{(\sigma(U)-1)}{q_i \sigma(U)} \left( \frac{q_i}{G_i(U)} \right)^{\frac{\sigma(U)-1}{\sigma(U)}} = \mu p_i$$

where $\mu$ is a constant term (combination of the Lagrange multiplier associated with the equation in $U$ and the budget constraint multiplier). To satisfy the budget constraint, $\frac{(\sigma(U)-1)}{\sigma(U)}$ has to equal $1/w$.

In other words, $\left( \frac{q_i}{G_i(U)} \right)^{\frac{\sigma(U)-1}{\sigma(U)}}$ corresponds to the budget share of good $i$ in consumption baskets:

$$\left( \frac{q_i}{G_i(U)} \right)^{\frac{\sigma(U)-1}{\sigma(U)}} = \frac{(\sigma(U)-1)\mu}{\sigma(U)} p_i q_i = \frac{p_i q_i}{w}.$$ 

This leads to the demand $q_i$:

$$q_i = G_i(U)^{1-\sigma(U)} (p_i/w)^{-\sigma(U)}$$

which is the same expression as above, with $\Lambda$ corresponding to utility. Moreover, we can see that utility $U$ is such that $\sum_i \left( \frac{q_i}{G_i(U)} \right)^{\frac{\sigma(U)-1}{\sigma(U)}} = 1$ which, using the demand for $q_i$ just above, can be written as:

$$\sum_i [G_i(U) p_i/w]^{1-\sigma(U)} = 1$$
which is the same equation as the one determining $\Lambda$, which proves that $\Lambda = U$.

**Proof of equivalence between condition ii) and inequality (23)** We mention in the text that condition ii) of Proposition 4 is equivalent to inequality (23) when both $\sigma$ and $G_i$ are differentiable.

Taking the derivative of the log of $G_i(\Lambda)\alpha_i^{-\frac{1}{1-\sigma(\Lambda)}}$ with respect to $\Lambda$, we find that it is positive if and only if:

$$\frac{G_i'(\Lambda)}{G_i(\Lambda)} - \log(\alpha_i) \cdot \frac{\partial}{\partial \Lambda} \left(\frac{1}{1-\sigma(\Lambda)}\right) > 0.$$ 

Hence, for each good $i$, the minimum $\alpha_i$ such that it is positive is:

$$\alpha_i^* = \exp\left(\frac{(\sigma(\Lambda) - 1)^2 G_i'(\Lambda)}{\sigma'(\Lambda) G_i(\Lambda)}\right).$$

One can see that inequality $\sum_i \alpha_i^* < 1$ corresponds to inequality (23) in the text.

Note: one can also verify that this condition is equivalent to imposing that $G_i(\Lambda)$ and $\sigma(\Lambda)$ are such that:

$$\left[\sum_i (G_i(\Lambda) p_i/w)^{1-\sigma(\Lambda)}\right]^{\frac{1}{1-\sigma(\Lambda)}}$$

increases for any set of $p_i/w$.

**Lemma 9** For any given set of $x_i \geq 0$ and $\alpha_i \geq 0$ such that $\sum_i \alpha_i = 1$, the following expression is monotonically increasing in $\rho \in (-\infty, +\infty)$:

$$\left[\sum_i \alpha_i x_i^\rho\right]^\frac{1}{\rho}$$

**Proof of Lemma 9:** First, consider two values $\rho < \rho' < 0$ and consider the mapping $m(x) = x^{\rho'}$ which is convex in $x$. Jensen’s inequality implies that:

$$m\left(\sum_i \alpha_i y_i\right) \leq \sum_i \alpha_i m(y_i)$$

and thus:

$$\left(\sum_i \alpha_i y_i\right)^\frac{1}{\rho} \leq \left(\sum_i \alpha_i y_i^{\rho'}\right)^\frac{1}{\rho'}.$$ 

Choosing $y_i = [x_i]^\rho$, we obtain:

$$\left[\sum_i \alpha_i x_i^\rho\right]^\frac{1}{\rho} \leq \left[\sum_i \alpha_i x_i^{\rho'}\right]^\frac{1}{\rho'}.$$ 

Second, consider two values $\rho' > \rho > 0$ and consider again the mapping $m(x) = x^{\rho'}$ which is now concave in $x$. Jensen’s inequality for concave functions implies:

$$m\left(\sum_i \alpha_i y_i\right) \geq \sum_i \alpha_i m(y_i)$$
and thus, taking to the exponent $1/\rho < 0$, we have:

$$\left( \sum_i \alpha_i y_i \right)^{\frac{1}{\rho}} \leq \left( \sum_i \alpha_i y_i^{\rho'} \right)^{\frac{1}{\rho'}}.$$ 

Choosing $y_i = [x_i]^{\rho}$, we obtain:

$$\left[ \sum_i \alpha_i x_i^{\rho} \right]^{\frac{1}{\rho}} \leq \left[ \sum_i \alpha_i x_i^{\rho'} \right]^{\frac{1}{\rho'}}.$$ 

Note that these terms are well defined when $\rho$ converges to zero (on both sides):

$$\lim_{\rho\to0} \left[ \sum_i \alpha_i x_i^{\rho} \right]^{\frac{1}{\rho}} = \prod_i x_i^{\alpha_i}$$

hence the findings above also apply to $\rho = 0$. This proves Lemma 9.

**Counter-examples when condition [A4] fails.**

Here I provide counter-examples to show that $\Lambda$ or $U$ are not well defined if the assumptions of Proposition 4 are not satisfied.

- First, suppose that $\sigma(\Lambda)$ increases in $\Lambda$. In this case, the elasticity of substitution increases with income and issues are more likely to arise when consumption is concentrated in one or few goods.

  When $G_i(\Lambda)$ is not monotonic in $\Lambda$ for a good $i$, the budget constraint can be written:

  $$G_i(\Lambda)p_i/w = 1$$

  when the consumption of all other goods become negligible, i.e. when $(p_j/w)^{1-\sigma(\Lambda)} = 0$. If there exists $\Lambda_1 \neq \Lambda_2$ such that $G_i(\Lambda_1) = G_i(\Lambda_2)$, one can see that the equation above has at least two solutions when $p_i/w = 1/G_i(\Lambda_1)$.

  Conversely, utility is not well defined by the implicit equation provided in Proposition 4 when $G_i$ is not monotonic for a good. Suppose that $q_j^{\sigma(U) - 1}$ is zero (or close to zero) for other goods $j$. In that case, we can see that $\left( q_i \frac{G_i(U)}{G_i(\Lambda)} \right)^{\frac{\sigma(U) - 1}{\sigma(U)}} = 1 \iff G_i(U) = q_i$ has several solutions in $U$ for some $q_i$ if $G_i$ is not monotonic, potentially violating the monotonicity of $U$ w.r.t. quantities.

  We also need $G_i'$ to have the same sign for all goods. If it is not the case, we can obtain situations where $\Lambda$ and $U$ are not well defined, or where $U$ would decrease with quantities $q_i$ for some goods.

- Counter-examples for the second case are more difficult to construct. Here we will assume here that $\sigma(\Lambda)$ and $G_i(\Lambda)$ are differentiable. Let us examine what happens when inequality (23) is not satisfied, i.e. when:

  $$\sum_i \exp \left( \frac{(\sigma(\Lambda) - 1)G_i'(\Lambda)}{\sigma'(\Lambda)G_i(\Lambda)} \right) > 1$$

  for a given $\Lambda = U_0$. In that case, we can show that it is possible to find a set of quantities $q_i$ such that $U_0$ is the solution of equation (24) but where implicit utility would depend negatively
on some of the quantities. This amounts to showing that the following expression:

\[
\left[ \sum_i \left( \frac{G_i(U)}{q_i} \right)^{1-\frac{\sigma(U)}{\sigma'(U)}} \right]^{\frac{\sigma(U)}{1-\sigma(U)}}
\]

decreases with \( U \) and for at least some of the \( q_i \)'s.

Suppose that \( U_0 \) is the solution of equation (24) for a given set of \( q_i \). We can always rearrange the \( q_i \) to match a given set of consumption shares while still having \( U_0 \) as the solution of equation (24). In particular, choose \( q_i^* \) such that \( U_0 \) is still the solution of (24) and such that:

\[
\left( \frac{G_i(U)}{q_i^*} \right)^{1-\frac{\sigma(U_0)}{\sigma'(U_0)}} = \frac{1}{A} \exp \left( \frac{(\sigma(U_0) - 1)^2G'_i(U_0)}{\sigma'(U_0)G_i(U_0)} \right)
\]

where \( A \equiv \sum_i \exp \left( \frac{(\sigma(U_0) - 1)^2G'_i(U_0)}{\sigma'(U_0)G_i(U_0)} \right) > 1 \), strictly larger than unity if condition ii) is not satisfied. Consider the function:

\[
f(U, q) = \left[ \sum_i \left( \frac{G_i(U)}{q_i} \right)^{1-\frac{\sigma(U)}{\sigma'(U)}} \right]^{\frac{\sigma(U)}{1-\sigma(U)}}
\]

which corresponds to the left-hand side of equation (24). One can see that the derivative in \( U \) at \( U = U_0 \) and \( q = q^* \) is negative:

\[
f_U(U_0, q^*) = \sum_i \frac{G'_i(U_0)}{G_i(U_0)} \left( \frac{G_i(U_0)}{q_i^*} \right)^{1-\frac{\sigma(U_0)}{\sigma'(U_0)}} + \frac{\sigma'(U_0)}{(1-\sigma(U_0))^2} \sum_i \left( \frac{G_i(U_0)}{q_i^*} \right)^{1-\frac{\sigma(U_0)}{\sigma'(U_0)}} \log \left( \frac{G_i(U_0)}{q_i^*} \right)^{1-\frac{\sigma(U_0)}{\sigma'(U_0)}}
\]

\[
= \frac{\sigma'(U_0)}{(1-\sigma(U_0))^2} \log A < 0
\]

while the derivative \( f_{q_i}(U_0, q^*) \) in each \( q_i \) is also negative. This leads to an implicit utility function \( U \) of \( q \) that decreases with quantities.

**Proof of Proposition 5**

**About \( \Lambda \).** Before we prove parts i) and ii) of Proposition 5, note that conditions [A5] ensure that \( \Lambda \) can be implicitly defined by the budget constraint as either a function of quantities \( q \) or normalized prices \( p/w \). Focusing on characterizing \( \Lambda \) as a function of quantities, we can follow the same approach as in Step 1 of the proof of Proposition 3.

First, using part ii) of restrictions [A5], we can see that the elasticity of \( D_i(F(\Lambda, U)y_i)/H(\Lambda, U) \) w.r.t. \( \Lambda \) is given by \( \varepsilon_F\varepsilon_D - \varepsilon_H \) which is assumed to be negative, hence it strictly decreases with \( \Lambda \). Symmetrically, we obtain that \( D_i^{-1}(H(\Lambda, U)q_i, U)/F(\Lambda, U) \) also strictly decreases with \( \Lambda \): its elasticity w.r.t. \( \Lambda \) is \( \varepsilon_H/\varepsilon_D - \varepsilon_F \), which is also negative given that \( \varepsilon_D \) is negative and \( \varepsilon_H - \varepsilon_D\varepsilon_F \) is positive. Note that the budget constraint can be written as:

\[
\sum_i q_i D_i^{-1}(q_i H(\Lambda, U), U) / F(\Lambda, U) = 1. \tag{75}
\]

Adding up across goods, we obtain that the left-hand side of equation (75) decreases strictly with \( \Lambda \). This implies that the solution in \( \Lambda \) to equation (26) is unique (if it exists).
Existence is then guaranteed using condition \([A5]-\text{iii})\), which we can symmetrically reformulate in terms of quantities. We assume that, for any good \(i\), \(y_i > 0\) and \(V\), there exists \(\Lambda \in \mathbb{R}\) such that:

\[ y_i D_i (y_i F(\Lambda, V), V) / H(\Lambda, V) = 1/N. \]

Denote by \(D_i^{-1}\) the inverse with respect to the first argument of \(D_i\). Note that this equality is equivalent to \(1/N = 1/(Ny_i)D_i^{-1}(1/(Ny_i)H(\Lambda, V), V) / F(\Lambda, V).\)

Hence, denoting \(q_i = 1/Ny_i\), we obtain that for any good \(i\), \(q_i > 0\) and \(U\), there exists \(\Lambda \in \mathbb{R}\) such that:

\[ q_i D_i^{-1}(\Lambda, U) / F(\Lambda, U) = 1/N. \]

For a given vector of quantities \(q\), for each good we obtain a \(\Lambda\) such that the equality above holds. Taking the maximum of the \(\Lambda\)'s obtained across the \(N\) goods (and using the monotonicity property in \(\Lambda\) described just above), we obtain a \(\Lambda_{\text{max}}\) such that

\[ \sum_i q_i D_i^{-1}(\Lambda_{\text{max}}, U) / F(\Lambda_{\text{max}}, U) \leq 1 \]

For the same vector \(q\), by taking the minimum of such \(\Lambda\)'s across goods, we obtain a \(\Lambda_{\text{min}}\) such that

\[ \sum_i q_i D_i^{-1}(\Lambda_{\text{min}}, U) / F(\Lambda_{\text{min}}, U) \geq 1. \]

As the left hand side of this expression is continuous in \(\Lambda\), the Intermediate Value Theorem ensures that a solution to equation (75) exists between \(\Lambda_{\text{min}}\) and \(\Lambda_{\text{max}}\).

Hence the budget constraint, i.e. equation (75), can be used to uniquely define \(\Lambda\) as a function of \(q\) and \(U\), or just as a function of \(q\) when we evaluate \(U\) at \(U(q)\). Moreover, as in Proposition 3, since the left-hand side of equation (75) has a strictly non-zero (negative) derivative in \(\Lambda\), we can use the Implicit Function Theorem to compute the derivatives of \(\Lambda\). Here, we now have an equation that also depend on \(U\), but we can still use the results as before (see Proposition 3) to compute the derivative \(\frac{\partial \Lambda}{\partial q_i}\) w.r.t. \(q_i\) for each goods \(i\) along an indifference curve, i.e. holding \(U\) constant. This partial derivative will be useful for the proof of quasi-concavity of \(U\), as discussed in part ii) further below.

i) Characterizing utility

The first part of Proposition 5 provides an equation that must be satisfied if demand can be rationalized and takes the form:

\[ q_i(p_i/w, \Lambda, V) = \frac{1}{H(\Lambda, V)} D_i \left( \frac{p_i F(\Lambda, V)}{w}, V \right) \]

or in terms of inverse demand:

\[ \frac{p_i}{w} = \frac{D_i^{-1}(H(\Lambda, U)q_i, U)}{F(\Lambda, U)} \]

satisfying equation (75). If demand can be rationalized with a differentiable utility function \(U(q)\), there exists a function \(\lambda(q)\) (real mapping from \(\mathbb{R}_+^N\) to \(\mathbb{R}_+\), such that

\[ \frac{\partial U}{\partial q_i} = \lambda(q) \frac{1}{H(\Lambda, U)} D_i \left( \frac{p_i F(\Lambda, U)}{w}, U \right). \]

Define a function \(M(q, \Lambda, U)\) as:

\[ M(q, \Lambda, U) = \sum_i \int_{q=q_0}^{q_i H(\Lambda, U)} D_i^{-1}(q, U) dq - \int_{\Lambda' = \Lambda_0}^{\Lambda} \frac{\partial H}{\partial \Lambda}(\Lambda', U) F(\Lambda', U) d\Lambda' \]
The partial derivative of $M$ w.r.t. $q$ is:

$$\frac{\partial M}{\partial q_j} = H(\Lambda, U)D_j^{-1}(q_j H(\Lambda, U), U)$$

The partial derivative of $M$ w.r.t. $\Lambda$ is:

$$\frac{\partial M}{\partial \Lambda} = \frac{\partial H}{\partial \Lambda}(\Lambda, U) \sum_i q_i D_i^{-1}(q_i H(\Lambda, U), U)dq - \frac{\partial H}{\partial \Lambda}(\Lambda, U)F(\Lambda, U)$$

Note that this partial derivative null at $\Lambda = \Lambda(q)$ and $U = U(q)$ if the budget constraint is satisfied (condition 75).

Now, define $\tilde{M}(q) = M(q, \Lambda(q), U(q))$, i.e. equal to $M$ where $U$ and $\Lambda$ are evaluated at $U(q)$ and $\Lambda(q)$ respectively rather than treated as arguments. Note that, if demand is rational, marginal utility must be itself proportional to inverse demand and thus:

$$H(\Lambda, U)D_j^{-1}(q_j H(\Lambda, U), U) = H(\Lambda, U)D_j^{-1}(q_j H(\Lambda, U), U) = \frac{H(\Lambda, U)F(\Lambda, U)}{\lambda(q)} \frac{\partial U}{\partial q_j}$$

where $\lambda$ is the Lagrange multiplier associated with the budget constraint, and where $U$ and $\Lambda$ are evaluated at $U(q)$ and $\Lambda(q)$. We obtain that the gradient of $\tilde{M}$ is proportional to the gradient of utility:

$$\frac{\partial \tilde{M}}{\partial q_j} = \frac{\partial M}{\partial q_j} + \frac{\partial M}{\partial \Lambda} \frac{\partial \Lambda}{\partial q_j} + \frac{\partial M}{\partial U} \frac{\partial U}{\partial q_j}$$

$$= H(\Lambda, U)D_j^{-1}(q_j H(\Lambda, U), U) + 0 + \frac{\partial M}{\partial U} \frac{\partial U}{\partial q_j}$$

$$= \left[ \frac{H(\Lambda, U)F(\Lambda, U)}{\lambda} + \frac{\partial M}{\partial U} \right] \frac{\partial U}{\partial q_j}.$$ 

Given that indifference curves are connected, this implies that there exist a function $\tilde{M}(U)$ of utility such that: $\tilde{M}(q) = \tilde{M}(U(q))$ for all $q$ (see e.g. Lemma 1 of Goldman and Uzawa, 1964, for a proof of this statement). Hence, combining with equation (76), we obtain that $U(q)$ satisfies:

$$\tilde{M}(U(q)) = \sum_i \int_{q=q_0}^{q_i H(\Lambda(q), U(q))} D_i^{-1}(q, U(q))dq - \int_{\Lambda' = \Lambda_0}^{\Lambda(q)} \frac{\partial H}{\partial \Lambda}(\Lambda', U(q))F(\Lambda', U(q))d\Lambda'. \quad (77)$$

Defining $G(\Lambda, U)$ as:

$$G(\Lambda, U) = \tilde{M}(U) + \int_{\Lambda' = \Lambda_0}^{\Lambda(q)} \frac{\partial H}{\partial \Lambda}(\Lambda', U(q))F(\Lambda', U(q))d\Lambda'.$$
As described in Proposition 5, we obtain that $U(q)$ must satisfy:

$$
\sum_i \int_{q=q_0}^{q_i} H(\Lambda(q), U(q)) D_i^{-1}(q, U(q)) dq - G(\Lambda(q), U(q)) = 0.
$$

(78)

**ii) Properties of the utility function**

We now examine the converse of part i): assuming that such equation admits a solution in $U$, does it yield a well-behaved utility function that is monotonic in each $q_i$, continuous and quasi-concave?

First, continuity is ensured by the fact that the left-hand side of equation (78) is continuous in $q$, $\Lambda$ and $U$, and is assumed to strictly decrease with $U$ (and $\Lambda$ is itself a differentiable function of $q$). Hence we can solve for $U$ as a continuous function of $q$.

Second, note that the left-hand side of equation (78) is strictly increasing in $q_i$, with a partial derivative (holding $U$ constant) given by:

$$
\frac{\partial M}{\partial q_j} = H D_i^{-1}(q_j, H, U) > 0
$$

(with the partial derivative in $\Lambda$ being null). As we assume that the left-hand side of equation (78) strictly decreases with $U$, the solution for $U$ must be strictly increasing in $q_i$ for each good $i$.

**Quasi-concavity of $U$.** Third and least obvious, we need to prove that the solution for utility $U$ is quasi-concave in $q$. To do so, we can however build up on Step 2 of the proof of Proposition 3 provided earlier. In Proposition 3, we have already shown the quasi-concavity of the following function $B$, holding $U$ constant:

$$
B(q, U) = \sum_i \int_{q=q_0}^{q_i} H(\Lambda^*(q, U), U) D_i^{-1}(q, U) dq - G(\Lambda^*(q, U), U)
$$

(function $B$ replaces the former utility function $U$ in Proposition 3) with $\Lambda^*(q, U)$ defined such that the following condition holds:

$$
\sum q_i D_i^{-1}(q_i, H(\Lambda(q), U), U) = F(\Lambda, U)
$$

again for a given $U$, where $F$ is such that $\frac{\partial G}{\partial \Lambda}(\Lambda, U) = \frac{\partial H}{\partial \Lambda}(\Lambda, U) F(\Lambda, U)$.

We can then use the quasi-concavity of $B$ (holding $U$ constant) to prove the quasi-concavity of $U$, defined implicitly by $B(q, U(q)) = 0$ (this implicit definition is equivalent to equation 78). The quasi-concavity of $B$ implies that for any $q$ and $q'$ such that $B(q, U) = B(q', U) = 0$, we must have:

$$
B(\alpha q + (1 - \alpha) q', U) \geq B(q, U) = 0.
$$

We can check also that the derivative of $B$ in $U$ is negative if $B(q, U) = 0$. Hence we obtain that: utility $U'$ evaluated at $(\alpha q + (1 - \alpha) q')$ is larger than $U(q) = U(q')$:

$$
U' \equiv U(\alpha q + (1 - \alpha) q') \geq U
$$

for any $\alpha \in (0, 1)$, since $B$ is strictly decreasing in $U$ and since $U''$ must satisfy:

$$
B(\alpha q + (1 - \alpha) q', U'') = 0.
$$

The fact that $U(\alpha q + (1 - \alpha) q') \geq U(q)$ whenever $U(q) = U(q')$ means that $U$ is quasi-concave. We can also check that $U$ is strictly quasi-concave if $B$ is quasi-concave.
Indirect utility for the two-aggregator case

As for the single-aggregator case, we obtain:

\[
\int_{q_0}^{H(\Lambda, U)q_i} D_i^{-1}(q, U) dq = - \int_{D_i^{-1}(q_0, U)}^{F(\Lambda, U)p_i/w} D_i(y, U) dy + D_i(F(\Lambda, U)p_i/w, U) F(\Lambda, U)p_i/w - D_i^{-1}(q_0, U)q_0
\]

which holds for a given level of utility \(U\). Moreover, note that we have:

\[
\sum_i (p_i/w, V) D_i(F(\Lambda, V)p_i/w, V) = H(\Lambda, V).
\]

Hence, summing across goods, we obtain:

\[
\sum_i \int_{q_0}^{H(\Lambda, U)q_i} D_i^{-1}(q, U) dq = - \sum_i \int_{D_i^{-1}(q_0, U)}^{F(\Lambda, U)p_i/w} D_i(y, U) dy + H(\Lambda, U)F(\Lambda, U) - \sum_i D_i^{-1}(q_0, U)q_0.
\]

Applying these equalities to the expression for direct utility provided in the text, we obtain a similar condition characterizing (indirect) utility as a function of normalized prices:

\[
\sum_i \int_{q_0}^{H(\Lambda, U)q_i} D_i^{-1}(q, U) dq = G(\Lambda, U)
\]

\[
\Leftrightarrow \sum_i \int_{D_i^{-1}(q_0, U)}^{F(\Lambda, U)p_i/w} D_i(y, U) dy = -G(\Lambda, U) + H(\Lambda, U)F(\Lambda, U) - \sum_i D_i^{-1}(q_0, U)q_0.
\]

Next, using our definition of \(G\) (using function \(\tilde{\tilde{M}}(U)\) defined above in the proof of Proposition 5) and integrating by parts, note that we have:

\[
G(\Lambda, U) = \tilde{\tilde{M}}(U) + \int_{\Lambda' = \Lambda_0}^{\Lambda} \frac{\partial H}{\partial \Lambda}(\Lambda', U)F(\Lambda', U)d\Lambda'
\]

\[
= \tilde{\tilde{M}}(U) + H(\Lambda, U)F(\Lambda, U) - H(\Lambda_0, U)F(\Lambda_0, U) - \int_{\Lambda' = \Lambda_0}^{\Lambda} \frac{\partial F}{\partial \Lambda}(\Lambda', U)H(\Lambda', U)d\Lambda'
\]

and thus the equality above is equivalent to:

\[
\sum_i \int_{q_0}^{F(\Lambda, V)p_i/w} D_i(y, V) dy = K(\Lambda, V)
\]

where function \(K\) is defined as:

\[
K(\Lambda, V) = \int_{\Lambda' = \Lambda_0}^{\Lambda} \frac{\partial F}{\partial \Lambda}(\Lambda', V)H(\Lambda', V)d\Lambda'
\]

\[
- \sum_i D_i^{-1}(q_0, V)q_0 - \tilde{\tilde{M}}(V) + H(\Lambda_0, V)F(\Lambda_0, V) + \sum_i \int_{q_0}^{D_i^{-1}(q_0, V)} D_i(y, V) dy.
\]

Notice that the second line only depends on \(V\), not \(\Lambda\), hence: \(\frac{\partial K}{\partial \Lambda}(\Lambda, V) = \frac{\partial F}{\partial \Lambda}(\Lambda, V)H(\Lambda, V)\).
Section 4) Practical cases and applications

Different forms of separability as special cases

Implicit separability If \( H \) does not depend on the aggregator \( \Lambda \), we have: \( \frac{\partial G}{\partial \Lambda}(\Lambda, U) = 0 \), hence \( G(\Lambda, U) = G(U) \). Without loss of generality, we can rescale function \( D_i \) by \( 1/G \) and impose \( G(U) = 1 \) after scaling. Utility \( U \) is then implicitly defined by

\[
\sum_i \int_{q=q_0}^{q_i} D_i^{-1}(q, U) \, dq = 1. 
\] (79)

In this case, \( F \) must be a monotonic function of the aggregator \( \Lambda \). It is also without loss of generality to assume \( F(\Lambda, U) = \Lambda \).

Then, if \( D_i^{-1}(q_i, U) \) is strictly decreasing in \( U \), and takes values from the full interval \((+\infty, 0)\) as \( U \) decreases (conditional on \( q_i \)), the utility function defined implicitly by this equation is uniquely defined, for any \( q_i \), and well-behaved.

Indirect implicit separability If \( F \) does not depend on the aggregator \( \Lambda \), we can rescale function \( D_i \) such that it is without loss of generality to assume that \( F = 1 \). This also implies that function \( K \) obtained in equation (28) only depends on \( V \), since \( \frac{\partial K}{\partial \Lambda}(\Lambda, V) = H(\Lambda, V) \frac{\partial F}{\partial \Lambda}(\Lambda, V) = 0 \). Hence, indirect utility can then be seen as the implicit solution of

\[
\sum_i \int_{y_i}^{p_i/w} D_i(y, V)/K(V) \, dy = 1. 
\]

Again, by rescaling \( D_i \) by \( K \), it is without loss of generality to assume \( K = 1 \).

If \( D_i(y, V) \) is strictly decreasing in \( V \), and takes values from the full interval \((+\infty, 0)\) as \( V \) decreases (conditional on \( q_i \)), the indirect utility function defined implicitly by this equation is uniquely defined for all sets of prices \( p/w \) and well-behaved.

Direct semi-separability Preferences as directly semi-separable if utility is:

\[
U(q) = \frac{1}{G(\Lambda)} \sum_i R_i(H(\Lambda)q_i) 
\] (80)

where \( H \), \( G \) and \( R_i \) are twice continuously-differentiable, with \( G' > 0 \), \( H' > 0 \), \( R_i' > 0 \) and \( R_i'' < 0 \) and where \( \Lambda \) is such that:

\[
\frac{\sum_i q_i R_i'(H(\Lambda)q_i)}{\sum_i R_i(H(\Lambda)q_i)} = \frac{F(\Lambda)}{G(\Lambda)} 
\] (81)

where \( F(\Lambda) \equiv G'(\Lambda)/H'(\Lambda) \).

This demand system is a special case of Proposition 5, as this corresponds to defining \( D_i(y_i, V) = R_i^{-1}(V y_i) \) and specifying \( F \), \( G \), and \( H \) as functions of \( \Lambda \) only. Here I provide again a derivation of demand for this special case.

First, note that the derivative of the right-hand-side of (80) is equal to:

\[
\frac{1}{G(\Lambda)^2} \left[ \sum_i q_i R_i'(H(\Lambda)q_i) H'(\Lambda) G(\Lambda) - \sum_i R_i(H(\Lambda)q_i) G'(\Lambda) \right] 
\]
which is null if condition (80) is satisfied. Hence marginal utility is given by the derivative of (80) holding $\Lambda$ constant. This yields:

$$\lambda p_i / w = \frac{\partial \tilde{U}}{\partial q_i} = \frac{H(\Lambda)}{G(\Lambda)} R'_i(H(\Lambda)q_i).$$

The budget constraint implies:

$$\lambda = \lambda \sum_i q_i p_i / w = \frac{H(\Lambda)}{G(\Lambda)} \sum_i q_i R'_i(H(\Lambda)q_i) = \frac{H(\Lambda)}{H'(\Lambda)G(\Lambda)} \sum_i R_i(H(\Lambda)q_i)G'(\Lambda) = \frac{H(\Lambda)G'(\Lambda)U}{H'(\Lambda)G(\Lambda)}$$

And thus we obtain the following expression for inverse demand:

$$p_i / w = \frac{H(\Lambda)R'_i(H(\Lambda)q_i)}{\lambda G(\Lambda)} = \frac{H'(\Lambda)R'_i(H(\Lambda)q_i)}{UG'(\Lambda)} = \frac{R'_i(H(\Lambda)q_i)}{UF(\Lambda)}$$

where $F(\Lambda) = G'(\Lambda)/H'(\Lambda)$. Re-inverting, we obtain Marshallian demand for good $i$:

$$q_i = R'^{-1}_i(V F(\Lambda)p_i / w) / H(\Lambda)$$

Conditions [A5]-ii) required by Proposition 5 is met if $R'^{-1}_i(F(\Lambda)y_i)/H(\Lambda)$ has a strictly negative derivative in $\Lambda$. Conditions iii) is met if this expression goes from $+\infty$ to 0 (in the limit) as $\Lambda$ increases. Hence equation (32) has a unique solution in the aggregator $\Lambda$.

Written as in Proposition 5, the condition characterizing utility is:

$$\sum_i R_i(H(\Lambda)q_i) / UG(\Lambda) = 1.$$ 

In this case, it is obvious that it is strictly decreasing in $U$ (holding $\Lambda$ and $q$ constant), and that a solution in $U$ exists.

**Indirect semi-separability** Preferences as indirectly semi-separable if indirect utility can be written:

$$V = \frac{\sum_i S_i(F(\Lambda)p_i / w)}{L(\Lambda)}$$

where $F$, $L$ and $S_i$ are twice continuously-differentiable, with $F' > 0$, $L' < 0$, $S'_i < 0$ and $S''_i > 0$, and where $\Lambda$ is such that:

$$\sum_i (p_i / w) D_i(F(\Lambda)p_i / w) / \sum_i S_i(F(\Lambda)p_i / w) = \frac{H(\Lambda)}{K(\Lambda)}$$

where we define $D_i(y_i) = -S'_i(y_i)$ and $H(\Lambda) = -L'(\Lambda)/F'(\Lambda)$.

Such indirect utility function is again a special case of the dual-aggregator form that we studied in Proposition 5, with $D_i(y_i, V) = -S'_i(y_i)/V$ and specifying $F$ and $H$ as functions of $\Lambda$ only. Condition [A5]-ii) required by Proposition 5 is met if $D_i(F(\Lambda)y_i)/H(\Lambda)$ has a strictly negative derivative in $\Lambda$. Condition [A5]-iii) is met if this term goes from $+\infty$ to 0 (in the limit) as $\Lambda$ increases.

Using Roy’s identity, we can check that demand for good $i$ equals:

$$q_i = \frac{D_i(F(\Lambda)p_i / w)}{V H(\Lambda)}.$$
We can switch for a characterization of indirect utility to a characterization of direct utility by integrating by part.

From equation (82), we obtain:

\[ \sum_i S_i(D_i^{-1}(UHq_i)) = LU. \]

From the budget constraint, we obtain:

\[ \sum_i (UHq_i)D_i^{-1}(UHq_i) = UHF. \]

Adding up the previous two equalities, we obtain:

\[ \sum_i S_i(D_i^{-1}(UHq_i)) + \sum_i (UHq_i)D_i^{-1}(UHq_i) = LU + UHF. \]

Denote by \( S_{0i} = \lim_{p \to +\infty} S_i(p) \), which is well defined since \( S_i \) is positive and decreasing. For each good \( i \), we have the following geometric equality (integration by part):

\[ S_i(D_i^{-1}(q)) + qD_i^{-1}(q) = S_{0i} + \int_0^q D^{-1}(q')dq'. \]

Plugging this into the previous equality and dividing by \( U \), we obtain:

\[ \sum_i \left[ \frac{S_{0i}}{U} + \frac{1}{U} \int_0^{UHq_i} D^{-1}(q')dq' \right] - (L + HF) = 0. \]

This equation in \( U \) corresponds to the characterization of utility in Proposition 5, with \( G(\Lambda) = L(\Lambda) + H(\Lambda)F(\Lambda) \). Note that the left-hand side is strictly decreasing in \( U \) so that the solution in \( U \) is unique.

**Symmetric homothetic QMOR**

Taking \( D_i(y) = \alpha_i y^{r-1} + \beta_i y^{\kappa-1} \) and \( F(\Lambda) = \Lambda \) and \( H(\Lambda) = \Lambda^{r-1} \), we obtain that the ideal price index \( P \) is then implicitly defined by:

\[ \sum_i \alpha_i \left( \frac{p_i\Lambda}{P} \right)^r + \sum_i \beta_i \left( \frac{p_i\Lambda}{P} \right)^{kr} - \Lambda^r = c_0. \]

for some constant term \( c_0 \), and where aggregator \( \Lambda \) satisfies:

\[ \sum_i \alpha_i \left( \frac{p_i\Lambda}{P} \right)^r + \sum_i \beta_i \left( \frac{p_i\Lambda}{P} \right)^{kr} = \Lambda^r. \]

Taking the difference between the previous two equations leads to:

\[ (P/\Lambda)^{kr} = \frac{1}{c_0} \left( \frac{1}{\kappa} - 1 \right) \sum_i \beta_i P_i^{kr}. \]
Normalizing $\frac{1}{\sigma_0} (\frac{1}{\kappa} - 1) = 1$ so that $\Lambda^{-\kappa r} = \sum_i \beta_i \left( \frac{P_i}{P} \right)^{\kappa r}$, we obtain a price index of such form:

$$P^r = \sum_i \alpha_i P_i^r + \left( \sum_i \beta_i P_i^{kr} \right)^{\frac{1}{\kappa}}.$$  

Taking the log derivative w.r.t. log price $p_i$, we obtain the expenditure share in good $i$ (Shepard’s Lemma):

$$\frac{p_i q_i}{w} = \alpha_i \left( \frac{P_i}{P} \right)^r + \beta_i \left( \frac{P_i}{P} \right)^{\kappa r} \Lambda^{-r(1-\kappa)}$$

and thus:

$$q_i = \frac{\alpha_i w}{P} \left( \frac{P_i}{P} \right)^{r-1} \left[ 1 + \frac{\beta_i}{\alpha_i} \left( \frac{\Lambda P_i}{P} \right)^{-r(1-\kappa)} \right] = \frac{w}{P} \left( \frac{P_i}{P} \right)^{r-1} \left[ \alpha_i + \beta_i P_i^{-r(1-\kappa)} \left( \sum_j \beta_j P_j^{\kappa r} \right)^{\frac{1}{\kappa}} \right].$$

With $\kappa = 1/2$, $\alpha_i = \alpha$ and $\beta_i = \beta$, we get symmetric QMOR used in Freenstra (2010). When $\alpha_i > 0$ and $\beta_i < 0$, note that we get a finite reservation price (choke price).

**A non-homothetic version of QMOR**

Here we adopt the notation from Mrázová and Neary (2013). The notation used previously for homothetic case corresponds to $r = 1 - \nu$ and $\kappa = (\sigma - 1)/(\nu - 1)$.

We have then:

$$\sum_i \alpha_i(V) \left( \frac{P_i}{w} \Lambda \right)^{1-\nu} + \frac{\nu - 1}{\sigma - 1} \sum_i \beta_i(V) \left( \frac{P_i}{w} \Lambda \right)^{1-\sigma} - \Lambda^{1-\nu} = c_0$$

where aggregator $\Lambda$ satisfies:

$$\sum_i \alpha_i(V) \left( \frac{P_i}{w} \Lambda \right)^{1-\nu} + \sum_i \beta_i(V) \left( \frac{P_i}{w} \Lambda \right)^{1-\sigma} - \Lambda^{1-\nu} = 0.$$ 

Taking the difference between the last two equations, we obtain:

$$\left( \frac{\nu - \sigma}{\sigma - 1} \right) \sum_i \beta_i(V) \left( \frac{P_i}{w} \Lambda \right)^{1-\sigma} = c_0.$$ 

Hence, setting $c_0 = \left( \frac{\nu - \sigma}{\sigma - 1} \right)$, we get:

$$\Lambda^{\sigma - 1} = \sum_i \beta_i(V) \left( \frac{P_i}{w} \right)^{1-\sigma}.$$ 

Plugging into the previous equation for $\Lambda$, we get:

$$\sum_i \alpha_i(V) \left( \frac{P_i}{w} \right)^{1-\nu} + \left( \sum_i \beta_i(V) \left( \frac{P_i}{w} \right)^{1-\sigma} \right)^{\frac{1-\nu}{\sigma - 1}} = 1.$$
Demand for good $i$ is then:

$$q_i = \alpha_i(V) \left( \frac{p_i}{w} \right)^{-\nu} + \beta_i(V) \Lambda^{\nu-\sigma} \left( \frac{p_i}{w} \right)^{-\sigma}.$$ 

We obtain the equation in the main text by plugging the expression for $\Lambda$.

**Linear demand**

Even with a simple linear demand in partial equilibrium, there are multiple ways to rationalize such demand functions with one or two aggregators.

Suppose that demand is linear for each good $i$ (with the caveat that preferences are satiated above a certain level). In the most general case with two aggregators $\Lambda$ and $V$, we obtain that demand must take the form:

$$q_i = \alpha_i(V) - F(\Lambda, V) \frac{p_i}{w} \frac{1}{\gamma_i(V)}$$

(or zero if the latter is negative), where $V$ is indirect utility and where $\Lambda$ satisfies:

$$\sum_i \left( \frac{p_i}{w} \right) \max \left\{ 0, \frac{\alpha_i(V) - F(\Lambda, V) \frac{p_i}{w}}{H(\Lambda, V) \gamma_i(V)} \right\} = 1$$

and which can be obtained from a utility that satisfies:

$$\sum_i \left[ \alpha_i(U) H(\Lambda, U) q_i - \frac{1}{2} \gamma_i(U) H(\Lambda, U)^2 q_i^2 \right] - G(\Lambda, U) = 0$$

where each $q_i$ must not exceed $\frac{\alpha_i(U)}{H(\Lambda, U) \gamma_i(U)}$. $\Lambda$ is uniquely defined if $H$ and $F$ are both increasing in $\Lambda$, and a solution in $\Lambda$ always exists if $H$ and $F$ span from 0 to $+\infty$ at the limit. In turn, the solution in $U$ is unique if we have the following monotonicity conditions (sufficient conditions), with strict monotonicity for at least one of them: $\alpha_i(U)$ decreases in $U$, $\gamma_i(U)$ increases in $U$, $H(\Lambda, U)$ decreases in $U$ and $G(\Lambda, U)$ increases in $U$.

To illustrate the versatility of this approach and the many ways to specify the demand shifters, several special cases are worth noting:

- Directly-additive preferences can generate such linear demand and yield: $q_i = \frac{\alpha_i - \Lambda p_i / w}{\gamma_i}$
- Indirectly-additive preferences yield: $q_i = \frac{\alpha_i - p_i / \Lambda}{\gamma_i}$
- Single aggregator preferences yields: $q_i = \frac{\Lambda \alpha_i - \Lambda^2 p_i / w}{\gamma_i}$
- Homothetic preferences yield: $q_i = \frac{w}{p} \cdot \frac{\alpha_i - F(\Lambda) p_i / P}{H(\Lambda) \gamma_i}$
- Directly implicitly-separable preferences yield: $q_i = \frac{\alpha_i(V) - \Lambda p_i / w}{\gamma_i(V)}$
- Indirectly implicitly-separable preferences yield: $q_i = \frac{\alpha_i(V) - p_i / w}{\Lambda \gamma_i}$
- Directly semi-separable preferences yield: $q_i = \frac{\alpha_i - F(\Lambda) V p_i / w}{H(\Lambda) \gamma_i}$
- Indirectly semi-separable preferences yield: $q_i = \frac{\alpha_i - F(\Lambda) p_i / w}{V H(\Lambda) \gamma_i}$
Translog cost function

Translog cost functions have been studied in a variety of contexts, from consumer theory to productivity estimation. While a general formulation specifies the price index as:

$$\log P = \alpha_0 + \sum_i \alpha_i \log p_i + \frac{1}{2} \sum_{i,j} \gamma_{ij} \log p_i \log p_j$$

with $\alpha_i > 0$, $\sum_i \alpha_i = 1$ and $\gamma_{ij} = \gamma_{ji}$ required for rationalization, applications often typically impose a symmetric parameterization across the $\gamma$’s, i.e. assume $\gamma_{ii} = \gamma/N - \gamma$ and $\gamma_{ij} = \gamma/N$ if $i \neq j$, with $\gamma > 0$.

As shown by Bergin and Feenstra (2009), the Symmetric Translog case leads to the following expenditure shares once we account for unavailable goods (or, equivalently, goods with prices above the choke price):

$$\frac{p_i q_i}{w} = \alpha_i + \left(1 - n \bar{\alpha}\right) + \gamma \left[\log p - \log p_i\right]$$

where $\log p$ denotes the average price across available varieties and $\bar{\alpha}$ is the average shifter $\alpha_i$ across available varieties, and $n$ is the number of available varieties with $q_i > 0$. Defining the aggregator as $\log \Lambda = -\log p - \left(1 - n \bar{\alpha}\right)/\gamma$, we can reformulate the expenditure share as:

$$\frac{p_i q_i}{w} = \alpha_i - \gamma \log(\Lambda p_i/w).$$

This corresponds to demand in Proposition 3 with $D_i(y) = \alpha_i - \gamma \log y$, $F(\Lambda) = 1/H(\Lambda) = \Lambda$, and is well defined even if such demand has a choke price. One can then notice that aggregator $\Lambda$ is uniquely determined by the budget constraint:

$$\sum_i \max \{0, \alpha_i - \gamma \log(\Lambda p_i/w)\} = 1$$

and that the price index can be obtained as:

$$\log P = \sum_i \alpha_i \log(\Lambda p_i/w) - \frac{\gamma}{2} \sum_i \left(\log(\Lambda p_i/w)\right)^2 - \log \Lambda.$$
where we define $K_i$ as:

$$K_i(q) = qD_i^{-1}(q).$$

Aggregators $Q$ and $F$ are initially defined as functions of vector of normalized prices, $p/w$. But since utility is assumed to be strictly quasi-concave, $p/w$ can be expressed as a function of the vector of quantities $q$. Hence $Q$ and $F$ can also be viewed as aggregators that are functions of quantities $q$, so that expenditure shares can be written as $W_i(q_j, Q(q), F(q)) = (Q(q)/F(q)) K_i(q_j/Q(q))$.

As stated, suppose that the set of gradients $\{\frac{\partial Q}{\partial \log p_i}, \frac{\partial F}{\partial \log p_i}\}$ is of rank two for all $(p, w)$. Invertibility of demand ($q$ as a function of $p/w$ and vice-versa) also ensures that the rank of $\{\frac{\partial Q}{\partial \log q_i}, \frac{\partial F}{\partial \log q_i}\}$ (as a function of normalized prices) is the same as the rank of $\{\frac{\partial Q}{\partial \log q_i}, \frac{\partial F}{\partial \log q_i}\}$ (as a function of quantities) evaluated at $q = q(p/w)$.

Differentiating the budget constraint $\sum_i K_i(q_i/Q) = F/Q$ implies:

$$\frac{\partial K_j}{\partial \log q_j} - \left(\sum_i \frac{\partial K_i}{\partial \log q_i}\right) \frac{\partial \log Q}{\partial \log q_j} = \frac{\partial (F/Q)}{\partial \log q_j}.$$ 

Hence, $\frac{\partial K_j}{\partial \log q_j}$ is colinear to the gradients $\frac{\partial Q}{\partial \log q_j}$ and $\frac{\partial F}{\partial \log q_j}$:

$$\frac{\partial K_j}{\partial \log q_j} = \frac{1}{Q} \frac{\partial F}{\partial \log q_j} + \left(\sum_i \frac{\partial K_i}{\partial \log q_i}\right) \frac{1}{Q} \frac{\partial Q}{\partial \log q_j} - \frac{F}{Q^2} \frac{\partial Q}{\partial \log q_j}. \quad (86)$$

If demand is rational and can be derived from utility maximization, we must have:

$$\frac{\partial U}{\partial \log q_i} = (\lambda Q/F) K_i(q_i/Q) \equiv \Lambda K_i(q_i/Q)$$

where we define the new aggregator $\Lambda = \lambda Q/F$ as a function of marginal utility $\lambda$ and the two aggregators $Q$ and $F$. Differentiating, we get:

$$\frac{\partial U}{\partial \log q_i \partial \log q_j} = \frac{\partial \Lambda}{\partial \log q_j} K_i - \Lambda \frac{\partial K_i}{\partial \log q_j} \frac{\partial \log Q}{\partial \log q_j}.$$

The cross derivative must be symmetric, hence, dividing by $\Lambda$ we obtain:

$$\frac{1}{\Lambda} \frac{\partial \Lambda}{\partial \log q_j} K_i - \frac{\partial K_i}{\partial \log q_j} \frac{\partial \log Q}{\partial \log q_j} = \frac{1}{\Lambda} \frac{\partial \Lambda}{\partial \log q_j} K_j - \frac{\partial K_j}{\partial \log q_j} \frac{\partial \log Q}{\partial \log q_j}.$$

Rearranging, and using again $\Lambda K_i = \frac{\partial U}{\partial \log q_i}$, we obtain:

$$\frac{1}{\Lambda^2} \frac{\partial \Lambda}{\partial \log q_j} \frac{\partial U}{\partial \log q_i} + \frac{\partial K_i}{\partial \log q_j} \frac{\partial \log Q}{\partial \log q_i} = \frac{1}{\Lambda^2} \frac{\partial \Lambda}{\partial \log q_j} \frac{\partial U}{\partial \log q_i} + \frac{\partial K_i}{\partial \log q_j} \frac{\partial \log Q}{\partial \log q_i}. \quad (87)$$

Incorporating (86) into (87) and simplifying, we obtain:

$$\frac{1}{\Lambda^2} \frac{\partial \Lambda}{\partial \log q_j} \frac{\partial U}{\partial \log q_i} + \frac{1}{Q^2} \frac{\partial F}{\partial \log q_j} \frac{\partial Q}{\partial \log q_i} = \frac{1}{\Lambda^2} \frac{\partial \Lambda}{\partial \log q_j} \frac{\partial U}{\partial \log q_i} + \frac{1}{Q^2} \frac{\partial F}{\partial \log q_j} \frac{\partial Q}{\partial \log q_i}. \quad (88)$$

The remainder of the proof exploits this symmetry condition (88) to show that $(\Lambda, U)$ can provide
an alternative set of aggregators to \((Q, F)\).

Take a vector \(x\) such that \(\sum_i x_i \frac{\partial \Lambda}{\partial \log q_i} = 0\). Multiplying equation (88) by \(x_i\) and summing across goods \(i\), we obtain:

\[
\frac{1}{\Lambda^2} \frac{\partial \Lambda}{\partial \log q_j} \left( \sum_i x_i \frac{\partial U}{\partial \log q_i} \right) + \frac{1}{Q^2} \frac{\partial F}{\partial \log q_j} \left( \sum_i x_i \frac{\partial Q}{\partial \log q_i} \right) = \frac{1}{Q^2} \left( \sum_i x_i \frac{\partial F}{\partial \log q_i} \right) \frac{\partial Q}{\partial \log q_j} \quad (89)
\]

If for all \(x\), we also get \(\sum_i x_i \frac{\partial U}{\partial \log q_i} = 0\), then we can see that the gradients of \(Q\) and \(F\) are colinear, which contradicts the assumption that they are not. Hence there exists \(x\) such that \(\sum_i x_i \frac{\partial U}{\partial \log q_i} \neq 0\) while we still have \(\sum_i x_i \frac{\partial \Lambda}{\partial \log q_i} = 0\). We can see from equation (89) that is implies that the gradient of \(U\) is colinear with the gradients of \(F\) and \(Q\).

Similarly, since the gradients of \(U\) and \(\Lambda\) are not colinear, we can find a vector \(z\) such that \(\sum_i z_i \frac{\partial U}{\partial \log q_i} = 0\) and \(\sum_i z_i \frac{\partial \Lambda}{\partial \log q_i} \neq 0\). Multiplying equation (88) by \(z_i\) and summing across goods \(i\), we obtain:

\[
\frac{1}{Q^2} \frac{\partial F}{\partial \log q_j} \left( \sum_i z_i \frac{\partial Q}{\partial \log q_i} \right) = \frac{1}{\Lambda^2} \frac{\partial \Lambda}{\partial \log q_j} \left( \sum_i z_i \frac{\partial U}{\partial \log q_i} \right) + \frac{1}{Q^2} \left( \sum_i z_i \frac{\partial F}{\partial \log q_i} \right) \frac{\partial Q}{\partial \log q_j} \quad (90)
\]

This implies that the gradient of \(\Lambda\) is also colinear with the gradients of \(F\) and \(Q\). Since the gradients of \(\Lambda\) and \(U\) are not colinear with each other, we obtain that the gradients of \(\Lambda\) and \(U\) offers an alternative basis on which we can project the gradients of \(F\) and \(Q\).

Aggregates \(F\) and \(Q\) can thus be written as functions of \(U(q)\) and such an aggregate \(\Lambda(q)\). Conversely, coming back to Marshallian demand instead of inverse demand, this also proves that we can express \(F\) and \(Q\) as a function of indirect utility \(V(p/w)\) and an aggregate \(\Lambda(p/W)\) that is function of normalized prices \(p/w\). Hence, such demand system is a special case of Proposition 2 and 5.

**Section 5) Application to monopolistic competition**

**Frechet differentiability with a continuum of goods**

The continuum of goods is \([0, \bar{N}]\), and a consumption profile is defined as \(q \in L^2[0, \bar{N}]\). From here onward, we denote the Lebesgue space \(L^n[0, \bar{N}]\) by \(L^n\) to simplify notation.

We would like to define utility implicit as a mapping from \(L^2\) to \(\mathbb{R}\) that satisfies:

\[
\int_{i=0}^{\bar{N}} \int_{q=0}^{\bar{N}} F(q_i U) D^{-1}(q, U) dq \, di \quad (91)
\]

where aggregator \(\Lambda\) is itself a solution to:

\[
\int_{i=0}^{\bar{N}} q_i D^{-1}(q_i H(\Lambda, U), U) \, di = 1 \quad (92)
\]

For utility to be well-defined and Frechet differentiable in \(q \in L^2\), the following conditions are needed:
First, note that the two integral sums in equations (91) and (92) are well-defined and finite for any \( q \in L^2 \). For the first one, we have:

\[
\int_{i=0}^{\bar{N}} \int_{q=0}^{H_q} D^{-1}(q,U) dq di < \int_{i=0}^{\bar{N}} \int_{q=0}^{HA} D^{-1}(q,U) dq di + D^{-1}(HA,U) \int_{i=0}^{\bar{N}} (q_i - A) 1_{\{q_i > A\}} di < +\infty
\]

for any constant term \( A > 0 \), since \( D^{-1}(q,U) \) is decreasing in \( q \). This integral is finite as we already assume that \( \int_{q=0}^{q_i} D^{-1}(q,U) dq \) is finite, and \( q \in L^2 \) (which implies that \( q \in L^1 \) since we are working over a bounded segment \([0, \bar{N}]\)). For the second one, note that we already assume \( \lim_{q_i \to 0} q_i D^{-1}(q_i,U) = 0 \) (i.e. the marginal utility form a good increases by less than \( 1/\gamma \) when \( q \) decreases).

\[
\int_{0}^{\bar{N}} q_i D^{-1}(q_i H, U) di < \int_{0}^{\bar{N}} q_i D^{-1}(q_i H, U) 1_{\{q_i > A\}} di + D^{-1}(AH,U) \int_{0}^{\bar{N}} q_i 1_{\{q_i > A\}} di < +\infty
\]

Next, as we define \( U \) implicitly as the solution of the system of equations (91) and (92), we need the Jacobian of the LHS to be well defined. The derivatives w.r.t. \( U \) depend on:

\[
\int_{i=0}^{\bar{N}} \int_{q=0}^{q_i} \frac{\partial D^{-1}}{\partial U} (q,U) dq di \quad \text{and} \quad \int_{i=0}^{\bar{N}} \frac{\partial D^{-1}}{\partial q} (q_i,U) di.
\]

We need to assume that those are well-defined and finite for any \( q \in L^2 \), a property that is not necessarily implied by the other assumptions made above.

The derivatives w.r.t. \( A \) are \( \int_{i=0}^{\bar{N}} q_i D^{-1}(q_i, U) di \) and \( \int_{i=0}^{\bar{N}} q_i \frac{\partial D^{-1}}{\partial q} (q_i, U) di \). The former one is finite, as shown above. The latter is finite if \( \left| \frac{\partial D^{-1}}{\partial q} \right| \) is bounded among large enough values of \( q \) and if it does not exceed \( A/q_i^2 \) for some constant term \( A \) in the limit \( q_i \to 0 \).

Note also that the Jacobian is triangular and invertible thanks to the assumptions that the derivative of LHS of equation (91) is strictly negative in \( U \), zero in \( \Lambda \) (this is implied by the budget constraint), and derivative of the LHS of equation (91) is strictly negative in \( \Lambda \).

Finally, for utility and \( \Lambda \) to be Fréchet differentiable, we need to assume that \( \int_{i=0}^{\bar{N}} \int_{q=0}^{q_i} D^{-1}(q,U) dq di \) and \( \int_{i=0}^{\bar{N}} q_i D^{-1}(q_i, U) di \) are Fréchet differentiable in \( q \). The derivatives are \( \int_{i=0}^{\bar{N}} D^{-1}(q_i, U) h_i di \) and \( \int_{i=0}^{\bar{N}} (D^{-1}(q_i, U) + q_i \frac{\partial D^{-1}}{\partial q}) h_i di \) respectively, for any \( h \in L^2 \). Hence Fréchet differentiability requires that:

\[
\int_{i=0}^{\bar{N}} \int_{q=0}^{q_i+h_i} D^{-1}(q,U) dq di - \int_{i=0}^{\bar{N}} D^{-1}(q_i,U) h_i di = o(||h||_2)
\]

and

\[
\int_{i=0}^{\bar{N}} (q_i+h_i) \left( D^{-1}(q_i+h_i, U) - D^{-1}(q_i, U) \right) di - \int_{i=0}^{\bar{N}} q_i \frac{\partial D^{-1}}{\partial q} h_i di = o(||h||_2)
\]

as \( h \) converges to zero, where \( || \cdot ||_2 \) denotes the \( L^2 \) norm.
Proof of Proposition 7

The proof relies on the results by Parenti et al. (2017), who provide simple conditions on the price elasticity (on the demand side) that determine the effect of population \( L \) and income \( w \) on key outcomes such as firm size and prices (markups). They express the price elasticity as a function of two variables to capture how consumers respond: average quantity \( q = Q/L \) consumed of each variety (\( x \) in their notation) and the number (measure) of firms \( N \). In particular, they show that:

i) an increase in population \( L \) leads to an increase in firm size and a decrease in prices if and only if \( E_q(\sigma) < E_N(\sigma) \), where \( E_q(\sigma) \) is the elasticity of the elasticity of substitution \( \sigma \) with respect to average sales \( q \) and \( E_N(\sigma) \) is the elasticity with respect to the number (measure) of firms \( N \).

ii) an increase in income \( w \) leads to an increase in firm size and a decrease in prices if and only if \( E_N(\sigma) > 0 \), i.e. if and only if the elasticity of substitution increases with the number of firms.

Taking stock of these results, the main task now is to reformulate these conditions with Gorman-Pollak demand with a single aggregator \( \Lambda \) as in Proposition 3.

First we examine how \( \Lambda \) depends on \( q \) and \( N \). Recall that \( \Lambda \) is implicitly determined by the budget constraint. With symmetry across all goods, the budget constraint (equation 53) becomes:

\[
(Nq) D^{-1}(H(\Lambda)q) = F(\Lambda)
\]

where \( q = Q/L \) is consumption by variety and by consumer. Taking the log derivative w.r.t. \( N \), we obtain:

\[
1 + (\varepsilon_H/\varepsilon_D) \frac{\partial \log \Lambda}{\partial \log N} = \varepsilon_F \frac{\partial \log \Lambda}{\partial \log N}
\]

which leads to:

\[
\frac{\partial \log \Lambda}{\partial \log N} = \frac{\varepsilon_D}{\varepsilon_F \varepsilon_D - \varepsilon_H}.
\]

Note that \( \varepsilon_F \varepsilon_D - \varepsilon_H \) and \( \varepsilon_D \) are both negative by assumption, hence \( \frac{\partial \log \Lambda}{\partial \log N} \) must be positive.

Taking now the log derivative of the budget constraint (93) w.r.t. \( q \), we obtain:

\[
(1 + 1/\varepsilon_D) + (\varepsilon_H/\varepsilon_D) \frac{\partial \log \Lambda}{\partial \log q} = \varepsilon_F \frac{\partial \log \Lambda}{\partial \log q}
\]

which leads to:

\[
\frac{\partial \log \Lambda}{\partial \log q} = \frac{\varepsilon_D + 1}{\varepsilon_F \varepsilon_D - \varepsilon_H}.
\]

Next, note that \( \sigma = -\varepsilon_D \) and is a function of \( H(\Lambda)q \). As in Mrazova and Neary (2014), we say that demand is superconvex if \( \sigma \) increases with \( H(\Lambda)q \) and subconvex if \( \sigma \) decreases with \( H(\Lambda)q \). Note that we take aggregators as given (partial equilibrium) to determine super or subconvexity.

In the superconvex case, one can then conclude that the price elasticity increases faster with \( N \) than with \( q \) (i.e. \( \mathcal{E}_N(\sigma) > \mathcal{E}_q(\sigma) \)) if and only if \( \varepsilon_H \frac{\partial \log \Lambda}{\partial \log N} > 1 + \varepsilon_H \frac{\partial \log \Lambda}{\partial \log q} \), while the opposite conclusion holds in the subconvex case. Plugging in the derivatives of \( \Lambda \) w.r.t. \( q \) and \( N \), we obtain that this statement is equivalent to:

\[
\varepsilon_H \frac{\partial \log \Lambda}{\partial \log N} > 1 + \varepsilon_H \frac{\partial \log \Lambda}{\partial \log q} \iff \frac{\varepsilon_H \varepsilon_D}{\varepsilon_F \varepsilon_D - \varepsilon_H} > 1 + \frac{\varepsilon_H \varepsilon_D + \varepsilon_H}{\varepsilon_F \varepsilon_D - \varepsilon_H} \iff 1 + \frac{\varepsilon_H \varepsilon_D}{\varepsilon_F \varepsilon_D - \varepsilon_H} < 0
\]
\[ \Leftrightarrow \frac{\varepsilon_H}{\varepsilon_H - \varepsilon_F \varepsilon_D} > 1 \]
\[ \Leftrightarrow \varepsilon_F < 0 \]

using the assumption that \( \varepsilon_D < 0 \) and \( \varepsilon_H - \varepsilon_F \varepsilon_D \) is positive (condition [A3]-i and [A3]-ii). This provides the condition for a positive impact of population on firm size (and negative impact on prices) in the superconvex case.

In the superconvex case, one can conclude that the price elasticity increases faster with \( N \) (i.e. \( \varepsilon N(\sigma) > 0 \)) if and only if \( \varepsilon_H \frac{\partial \log \Lambda}{\partial \log N} > 0 \), while the opposite conclusion holds in the subconvex case. Plugging in the derivatives of \( \Lambda \) w.r.t. \( N \), we obtain that this statement is equivalent to:

\[ \varepsilon_H \frac{\partial \log \Lambda}{\partial \log N} > 0 \Leftrightarrow \frac{\varepsilon_H \varepsilon_D}{\varepsilon_F \varepsilon_D - \varepsilon_H} > 0 \]
\[ \Leftrightarrow \varepsilon_H > 0 \]

using the assumption that \( \varepsilon_D < 0 \) and \( \varepsilon_H - \varepsilon_F \varepsilon_D > 0 \) (condition [A3]-i and [A3]-ii). This provides the condition for a positive impact of income on firm size (and negative impact on prices) in the superconvex case.

With the subconvex case associated with opposite conditions, we obtain Proposition 7:

i) an increase in population \( L \) leads to an increase in firm size and a decrease in prices iff \( \varepsilon_F < 0 \) in the superconvex case or \( \varepsilon_F > 0 \) in the subconvex case;

ii) an increase in income \( w \) leads to an increase in firm size and a decrease in prices iff \( \varepsilon_H > 0 \) in the superconvex case or \( \varepsilon_H < 0 \) in the subconvex case.

**With implicitly-additive preferences**

With directly-implicitly-additive preferences based on equation (29), and with symmetric demand over a continuum of goods, utility satisfies:

\[ N \int_0^q D^{-1}(q',U) dq' = 1 \]

where \( D^{-1} \) is strictly decreasing in \( q \) and \( U \). This implicitly determines how utility \( U \) depends on \( N \) and \( q \).

Market size effects are then determined by how \( \sigma = \frac{D^{-1}(q,U)}{q(D^{-1}(q,U))} \) depends on \( q \) and \( N \). Relying again on the results by Parenti et al. (2017), the price elasticity \( \sigma \) determines the effect of population \( L \) and income \( w \) on key outcomes such as firm size and prices (markups). In particular, they show that an increase in income \( w \) leads to an increase in firm size and a decrease in prices if and only if \( \mathcal{E}_N(\sigma) > 0 \) (see proof of Proposition 7 above), while directly-additive preferences impose \( \mathcal{E}_N(\sigma) = 0 \).

Here with implicitly-additive preferences, \( \sigma \) and \( D \) can be quite flexible, and can also yield both signs in \( \mathcal{E}_N(\sigma) \). For instance, take the preferences in Proposition 4: those are implicitly additive, with an elasticity of substitution \( \sigma \) that is solely a function of utility \( U \), and can decrease or increase with \( U \). This implies that \( \sigma \) can increase or decrease with \( N \) (as \( U \) increases with \( N \)), with both cases possible in theory. For instance, if \( \sigma \) decreases with utility \( U \), an increase in income \( w \) leads to a decrease in firm size and a more then proportional increase in the number of firms. The opposite holds if \( \sigma \) increases with \( U \).