

# Single Market Nonparametric Identification of Multi-Attribute Hedonic Equilibrium Models

Victor Chernozhukov, Alfred Galichon, Marc Henry, Brendan Pass

► **To cite this version:**

Victor Chernozhukov, Alfred Galichon, Marc Henry, Brendan Pass. Single Market Nonparametric Identification of Multi-Attribute Hedonic Equilibrium Models. 2014. <hal-01169655>

**HAL Id: hal-01169655**

**<https://hal-sciencespo.archives-ouvertes.fr/hal-01169655>**

Submitted on 29 Jun 2015

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# SINGLE MARKET NONPARAMETRIC IDENTIFICATION OF MULTI-ATTRIBUTE HEDONIC EQUILIBRIUM MODELS

VICTOR CHERNOZHUKOV, ALFRED GALICHON, MARC HENRY AND BRENDAN PASS

**ABSTRACT.** This paper derives conditions under which preferences and technology are nonparametrically identified in hedonic equilibrium models, where products are differentiated along more than one dimension and agents are characterized by several dimensions of unobserved heterogeneity. With products differentiated along a quality index and agents characterized by scalar unobserved heterogeneity, single crossing conditions on preferences and technology provide identifying restrictions. We develop similar shape restrictions in the multi-attribute case and we provide identification results from the observation of a single market. We thereby extend identification results in Matzkin (2003) and Heckman, Matzkin, and Nesheim (2010) to accommodate multiple dimensions of unobserved heterogeneity.

**Keywords:** Hedonic equilibrium, nonparametric identification, multidimensional unobserved heterogeneity, optimal transport.

**JEL subject classification:** C14, C61, C78

## 1. INTRODUCTION

Recent years have seen renewed interest in hedonic models, particularly their identification and estimation. Hedonic models were initially introduced to analyze price

---

*Date:* The present version is of December 29, 2014. The authors thank Guillaume Carlier, Pierre-André Chiappori and conference participants at the Fields Institute, Oberwolfach, various meetings of the Econometric Society and the Canadian Econometrics Study Group for stimulating discussions. Chernozhukov's research has received funding from the NSF. Galichon's research has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013), ERC grant agreement no 313699, and from FiME, Laboratoire de Finance des Marchés de l'Energie. Henry's research has received funding from SSHRC Grants 435-2013-0292 and NSERC Grant 356491-2013.

responses to quality parameters of differentiated goods. Among these: (1) Given the fact that the amenities offered by cars constantly evolve over time, how can one construct a price index measuring the evolution of car prices? (2) How can one explain price differentiation in wine, art, luxury goods, professional sports wages? (3) What does the correlation between the wage differentials and the level of risk associated to a given job reveal about individuals' valuation for their own life? (4) How can one analyze individual preferences for environmental features?

These questions gave rise to a vast literature, which aims at modeling implicit markets for quality differentiated products. There are two layers to this literature. The first layer is the literature on "hedonic regressions," which aims at estimating consumer willingness to pay for quality, while correcting for the standard endogeneity issue that consumers with greater taste for quality will consume more of it. The second layer, which concerns us here, has broader scope: the literature on "hedonic equilibrium models" incorporates a supply side with differentiated productivity over various quality parameters and studies the resulting equilibrium. This approach dates back at least as far as Tinbergen (1956); and Rosen (1974) provides a famous two-step procedure to estimate general hedonic models and thereby analyze general equilibrium effects of changes in buyer-seller compositions, preferences and technology on qualities traded at equilibrium and their price. Following the influential criticism of Rosen's strategy in Brown and Rosen (1982) and the inadmissibility of supply side observable characteristics as instruments in structural demand estimation as discussed in Epple (1987) and Bartik (1987), it was generally believed that identification in hedonic equilibrium models required data from multiple markets, as in Epple (1987), Khan and Lang (1988) and, more recently, Bajari and Benkard (2005) and Bishop and Timmins (2011).

Ekeland, Heckman, and Nesheim (2004) show, however, that hedonic equilibrium models are in fact identified from single market data, under separability assumptions, as in Ekeland, Heckman, and Nesheim (2004), or shape restrictions, as in Heckman,

Matzkin, and Nesheim (2010). The common underlying framework is that of a perfectly competitive market with heterogeneous buyers and sellers and traded product quality bundles and prices that arise endogeneously in equilibrium. Preferences are quasi-linear in price and under mild semicontinuity assumptions, Ekeland (2010) and Chiappori, McCann, and Nesheim (2010) show that equilibria exist, in the form of a joint distribution of product and consumer types (who consumes what), a joint distribution of product and producer types (who produces what) and a price schedule such that markets clear for each endogenously traded product type. Equilibrium existence results are valid in hedonic markets for multi-attribute products, but existing single market identification strategies restrict attention to a single quality dimension and scalar unobserved heterogeneity in consumer preferences and production technology. Ekeland, Heckman, and Nesheim (2004) require marginal utility (resp. marginal product) to be additively separable in unobserved consumer (resp. producer) characteristic. Heckman, Matzkin, and Nesheim (2010) show that demand is nonparametrically identified under a single crossing condition and that various additional shape restrictions allow identification of preferences without additive separability.

The objective of these papers and ours is to recover structural preference and technology parameters from the observation of who trades what and at what price. In the identification exercise, price is assumed known, as are the distributions characterizing who produces and consumes which good. Since price is observed and the environment is perfectly competitive, identification of preferences and identification of technology can be treated independently and symmetrically. Take the consumer problem, for instance. Under a single crossing condition on the utility function (also known as Spence-Mirlees in the mechanism design literature), the first order condition of the consumer problem yields an increasing demand function, i.e., quality demanded by the consumer as an increasing function of her unobserved type, interpreted as unobserved taste for quality. Assortative matching guarantees uniqueness of demand, as the unique increasing function that maps the distribution of unobserved taste for

quality, which is specified a priori, and the distribution of qualities, which is observed. Hence demand is identified as a quantile function, as in Matzkin (2003). Identification, therefore, is driven by a shape restriction on the utility function. We show that similar shape restrictions on the utility function also yield identification conditions in the case of non scalar characteristics and unobserved heterogeneity. In the special case, where marginal utility is additively separable in the unobservable taste vector, concavity yields nonparametric identification of the utility function, according to the celebrated Brenier Theorem of optimal transport theory (Theorem 3.8 of Villani (2003)). More generally, a generalization of single crossing known as the Twist Condition in optimal transport theory and a generalized convexity shape restriction yield identification of the utility function in hedonic equilibrium models with multiple quality dimensions. The distribution of unobserved heterogeneity is fully specified a priori and cannot be identified from single market data without additional separability conditions or exclusion restrictions.

*Related work.* Beyond Ekeland, Heckman, and Nesheim (2004), Heckman, Matzkin, and Nesheim (2010) and other contributions cited so far, this paper is closely related to the growing literature on identification and estimation of nonlinear econometric models with multivariate unobserved heterogeneity on the one hand, and to the empirical literature on matching models where agents match along multiple dimensions on the other hand. The quantile identification strategy of Matzkin (2003) was recently extended to non scalar unobserved heterogeneity using the Rosenblatt (1952)-Knothe (1957) sequential multivariate quantile transform for nonlinear simultaneous equations models in Matzkin (2013) and bivariate hedonic models in Nesheim (2013). Chiappori, McCann, and Nesheim (2010) derive a matching formulation of hedonic models and thereby highlight the close relation between empirical strategies in matching markets and in hedonic markets. Galichon and Salanié (2012) extend the work of Choo and Siow (2006) and identify preferences in marriage markets, where agents

match on discrete characteristics, as the unique solution of a programming problem, as in the present paper. The strategy is also applied in Chiong, Galichon, and Shum (2013) for identification of dynamic discrete choice problems and in Dupuy, Galichon, and Henry (2014) to discrete hedonic models.

*Organization of the paper.* The remainder of the paper is organized as follows. Section 2 sets the hedonic equilibrium framework out. Section 3 gives an account of the main results on nonparametric identification of preferences in single attribute hedonic models, mostly drawn from Ekeland, Heckman, and Nesheim (2004) and Heckman, Matzkin, and Nesheim (2010). Section 4 is the main section of the paper and shows how these results and the shape restrictions that drive them can be extended to the case of multiple attribute hedonic equilibrium markets. The last section discusses future research directions. Proofs of the main results are relegated to the appendix, as are necessary background results on optimal transport theory and hedonic equilibrium theory.

## 2. HEDONIC EQUILIBRIUM AND THE IDENTIFICATION PROBLEM

We consider a competitive environment, where consumers and producers trade a good or contract, fully characterized by its type or quality  $z$ . The set of feasible qualities  $Z \subseteq \mathbb{R}^{d_z}$  is assumed open and given a priori, but the qualities actually traded arise endogenously in the hedonic market equilibrium, as does their price schedule  $p(z)$ . Producers are characterized by their type  $\tilde{y} \in \tilde{Y} \subseteq \mathbb{R}^{d_{\tilde{y}}}$  and consumers by their type  $\tilde{x} \in \tilde{X} \subseteq \mathbb{R}^{d_{\tilde{x}}}$ . Type distributions  $P_{\tilde{x}}$  on  $\tilde{X}$  and  $P_{\tilde{y}}$  on  $\tilde{Y}$  are given exogenously, so that entry and exit are not modelled, except in the possibility of non participation, which is modelled by adding isolated points to the sets of types and renormalizing distributions accordingly (see Section 1.1 of Chiappori, McCann, and Nesheim (2010) for details). Consumers and producers are price takers and maximize quasi-linear utility  $U(\tilde{x}, z) - p(z)$  and profit  $p(z) - C(\tilde{y}, z)$  respectively. Utility  $U(\tilde{x}, z)$

(respectively cost  $C(\tilde{y}, z)$ ) is upper (respectively lower) semicontinuous and bounded and normalized to zero in case of nonparticipation. In addition, the set of qualities  $Z(\tilde{x}, \tilde{y})$  that maximize the surplus  $U(\tilde{x}, z) - C(\tilde{y}, z)$  for each pair of types  $(\tilde{x}, \tilde{y})$  is assumed to have a measurable selection. Then, Ekeland (2010) and Chiappori, McCann, and Nesheim (2010) show that an equilibrium exists in this market, in the form of a price function  $p$  on  $Z$ , a joint distribution  $P_{\tilde{x}, z}$  on  $\tilde{X} \times Z$  and  $P_{\tilde{y}, z}$  on  $\tilde{Y} \times Z$  such that their marginal on  $Z$  coincide, so that market clears for each traded quality  $z \in Z$ . Uniqueness is not guaranteed, in particular prices are not uniquely defined for non traded quantities in equilibrium. Purity is not guaranteed either: an equilibrium specifies a conditional distribution  $P_{z|\tilde{x}}$  (respectively  $P_{z|\tilde{y}}$ ) of qualities consumed by type  $\tilde{x}$  consumers (respectively produced by type  $\tilde{y}$  producers). The quality traded by a given producer-consumer pair  $(\tilde{x}, \tilde{y})$  is not uniquely determined at equilibrium.

Ekeland (2010) and Chiappori, McCann, and Nesheim (2010) further show that a pure equilibrium exists and is unique, under the additional assumption that type distributions  $P_{\tilde{x}}$  and  $P_{\tilde{y}}$  are absolutely continuous and gradients of utility and cost,  $\nabla_{\tilde{x}}U(\tilde{x}, z)$  and  $\nabla_{\tilde{y}}C(\tilde{y}, z)$  exist and are injective as functions of quality  $z$ . The latter condition, also known as the Twist condition in the optimal transport literature, ensures that a consumer of a given type  $\tilde{x}$  (respectively producer of a given type  $\tilde{y}$ ) will always consume (respectively produce) the same quality  $z$  at equilibrium.

The identification problem consists in the recovery of structural features of preferences and technology from observation of traded quantities and their prices in a single market. The solution concept we impose in our identification analysis is the following feature of hedonic equilibrium, i.e., maximization of surplus generated by a trade.

**Assumption EC.** [Equilibrium concept] The pair  $(\gamma, p)$ , where  $\gamma$  is a probability measure on  $\tilde{X} \times Z \times \tilde{Y} \subseteq \mathbb{R}^{d_{\tilde{x}}} \times \mathbb{R}^{d_z} \times \mathbb{R}^{d_{\tilde{y}}}$  and  $p$  is a function on  $Z$ , is an hedonic

equilibrium. That is:  $\gamma$  has marginals  $P_{\tilde{x}}$  and  $P_{\tilde{y}}$  and for  $\gamma$ -almost all  $(\tilde{x}, z, \tilde{y})$ ,

$$\begin{aligned} U(\tilde{x}, z) - p(z) &= \max_{z' \in Z} (U(\tilde{x}, z') - p(z')), \\ p(z) - C(\tilde{y}, z) &= \max_{z' \in Z} (p(z') - C(\tilde{y}, z')). \end{aligned}$$

In addition, observed qualities  $z = Z(\tilde{x}, \tilde{y})$  maximizing joint surplus  $U(\tilde{x}, z) - C(\tilde{y}, z)$  for each  $\tilde{x} \in \tilde{X}$  and  $\tilde{y} \in \tilde{Y}$ , lie in the interior of the set of feasible qualities  $Z$ .

Given observability of prices and the fact that producer type  $\tilde{y}$  (respectively consumer type  $\tilde{x}$ ) does not enter into the utility function  $U(\tilde{x}, z)$  (respectively cost function  $C(\tilde{y}, z)$ ) directly, we may consider the consumer and producer problems separately and symmetrically. We focus on the consumer problem and on identification of utility function  $U(\tilde{x}, z)$ . Under assumptions ensuring purity and uniqueness of equilibrium, the model predicts a deterministic choice of quality  $z$  for a given consumer type  $\tilde{x}$ . Hence, to preclude outright rejection of the model with any data set, where identical observable types consume different goods, we assume, as is customary, that consumer types  $\tilde{x}$  are only partially observable to the analyst. We write  $\tilde{x} = (x, \varepsilon)$ , where  $x \in \mathbb{R}^{d_x}$  is the observable part of the type vector, and  $\varepsilon \in \mathbb{R}^{d_\varepsilon}$  is the unobservable part. The observable and unobservable types will be assumed independent and we shall make a separability assumption that will allow us to specify constraints on the interaction between consumer unobservable type  $\varepsilon$  and good quality  $z$  in order to identify interactions between observable type  $x$  and good quality  $z$ .

**Assumption H.** [Unobservable heterogeneity] Consumer type  $\tilde{x}$  is composed of observable type  $x$  with distribution  $P_x$  on  $\mathbb{R}^{d_x}$  and unobservable type  $\varepsilon$  with a priori specified distribution  $P_\varepsilon$  on  $\mathbb{R}^{d_\varepsilon}$ . Observable and unobservable types are independent and the utility of consumers can be decomposed as  $U(\tilde{x}, z) = \overline{U}(x, z) + \zeta(x, \varepsilon, z)$ .

Assumption H includes independence of observable and unobservable types, which, although a strong requirement, is necessary in identification strategies that extend



the quantile identification of Matzkin (2003). Specification of  $P_\varepsilon$  is a necessary normalization, which also stems from quantile identification. Note that the separability assumption added here is vacuous until restrictions on the function  $\zeta$  are specified.

We shall work throughout under the following regularity condition on the primitives of the hedonic equilibrium model.

**Assumption R.** [Regularity of preferences and technology] The following regularity conditions hold.

- (1)  $U(\tilde{x}, z)$  and  $C(\tilde{y}, z)$  are differentiable with respect to  $z$ , for all  $\tilde{x}$  and  $\tilde{y}$ .
- (2)  $\zeta(x, \varepsilon, z)$  is continuously differentiable with respect to  $z$ , for all  $x, \varepsilon$  and  $z$ .
- (3)  $\det(\nabla_{z\varepsilon}^2 \zeta(x, \varepsilon, z)) \neq 0$  for all  $x, \varepsilon, z$ , where  $\nabla_{z\varepsilon}^2 \zeta = (\partial^2 \zeta / \partial \varepsilon_i \partial z)_{ij}$  is the  $n \times n$  matrix of mixed partial derivatives.

The object of inference is the deterministic component of utility  $\bar{U}(x, z)$ . We shall denote

$$V(x, z) = p(z) - \bar{U}(x, z)$$

and focus on identification of the function  $V$ , since under observability of price, it is equivalent to identification of the deterministic part of utility. We shall identify  $\bar{U}(x, z)$  up to a constant and use the following terminology.

**Definition 1** (Nonparametric identification). *The function  $\bar{U}(x, z)$  will be called non-parametrically identified if  $\nabla_z \bar{U}(x, z)$  is unique  $P_{z|x}$ -a.s. for all  $x$ .*

We shall work in stages, recalling first existing identification results in case of scalar  $z$  and clarifying which features we intend to extend and how. The guiding principle will be the characterization of shape restrictions on the function  $V$  that emulate single crossing and monotonicity restrictions in the scalar case and remain just identifying in the multi-attribute case.

## 3. SINGLE MARKET IDENTIFICATION WITH SCALAR ATTRIBUTE

We first recall and reformulate results of Heckman, Matzkin, and Nesheim (2010) on identification of single attribute hedonic models. Suppose, for the purpose of this section, that  $d_\varepsilon = d_z = 1$ , so that unobserved heterogeneity is scalar, as is the quality dimension. Suppose further (for ease of exposition) that  $\zeta$  is twice continuously differentiable in  $z$  and  $\varepsilon$  and that  $V$  is twice continuously differentiable in  $z$ . Consumers take price schedule  $p(z)$  as given and choose quality  $z$  to maximize  $\zeta(x, \varepsilon, z) - V(x, z)$ . We impose a single crossing condition on  $\zeta$ .

**Assumption S1.** [Spence-Mirlees] Dimension  $d_z$  of  $Z$  is 1 and for all  $x, \varepsilon, z$  of the domain of  $P_{\tilde{x}, z}$ ,  $\zeta_{\varepsilon z}(x, \varepsilon, z) > 0$ .

The first order condition of the consumer problem yields

$$\zeta_z(x, \varepsilon, z) = V_z(x, z), \quad (1)$$

which, under Assumption S1, implicitly defines an inverse demand function  $z \mapsto \varepsilon(x, z)$ , which specifies which unobserved type consumes quality  $z$ . Combining the second order condition  $\zeta_{zz}(x, \varepsilon, z) < V_{zz}(x, z)$  and further differentiation of (1), i.e.,  $\zeta_{zz}(x, \varepsilon, z) + \zeta_{\varepsilon z}(x, \varepsilon, z)\varepsilon_z(x, z) = V_{zz}(x, z)$ , yields

$$\varepsilon_z(x, z) = \frac{V_{zz}(x, z) - \zeta_{zz}(x, \varepsilon, z)}{\zeta_{\varepsilon z}(x, \varepsilon, z)} > 0.$$

Hence the inverse demand is increasing and is therefore identified as the unique increasing function that maps the distribution  $P_{z|x}$  to the distribution  $P_\varepsilon$ , namely the quantile transform. Denoting  $F$  the cumulative distribution function corresponding to the distribution  $P$ , we therefore have identification of inverse demand according to the strategy put forward in Matzkin (2003) as:

$$\varepsilon(x, z) = F_\varepsilon^{-1}(F_{z|x}(z|x)).$$

The single crossing condition of Assumption S1 on the consumer surplus function  $\zeta(x, \varepsilon, z)$  yields positive assortative matching, as in the Becker (1973) classical model. Consumers with higher taste for quality  $\varepsilon$  will choose higher qualities in equilibrium and positive assortative matching drives identification of demand for quality. The important feature of Assumption S1 is injectivity of  $\zeta_z(x, \varepsilon, z)$  relative to  $\varepsilon$  and a similar argument would have carried through under  $\zeta_{z\varepsilon}(x, \varepsilon, z) < 0$ , yielding negative assortative matching instead.

Once inverse demand is identified, the function  $V(x, z)$ , hence the utility function  $\bar{U}(x, z) = p(z) - V(x, z)$ , can be recovered up to a constant by integration of the first order condition (1):

$$V(x, z) = \int_0^z \zeta_z(x, \varepsilon(x, z'), z') dz'.$$

We summarize the previous discussion in the following identification statement, originally due to Heckman, Matzkin, and Nesheim (2010).

**Proposition 1.** *Under Assumptions EC, H, R and S1,  $\bar{U}(x, z)$  is nonparametrically identified.*

Unlike the demand function, which is identified without knowledge of the surplus function  $\zeta$ , as long as the latter satisfies single crossing (Assumption S1), identification of the preference function  $\bar{U}(x, z)$  does require a priori knowledge of the function  $\zeta$ . This includes existing results in this literature. Ekeland, Heckman, and Nesheim (2004) cover the special case, where  $\zeta(x, \varepsilon, z) = z\varepsilon$ . In that case,  $\varepsilon(x, z)$  increasing in  $z$ , maximizes, by the classical Hardy, Littlewood, and Pólya (1952) inequalities,  $\mathbb{E}[z\varepsilon|x]$  among all joint distributions for  $(z, \varepsilon)$ , subject to the marginal restrictions that  $z \sim P_{z|x}$  and  $\varepsilon \sim P_\varepsilon$ . We thereby recover the efficiency property of positive assortative matching and unlike positive assortative matching, the maximization of  $\mathbb{E}[z\varepsilon|x]$  has a natural extension in the multivariate case  $\zeta(x, \varepsilon, z) = z'\varepsilon$ , where quality  $z$  and taste for quality  $\varepsilon$  are conformable vectors. We shall examine the case  $\zeta(x, \varepsilon, z) = z'\varepsilon$  in the

next section, before moving to the general extension with arbitrary surplus function  $\zeta(x, \varepsilon, z)$ , allowing marginal utility to be nonlinear in unobserved taste, hence allowing interactions between consumer and good characteristics in the utility.

#### 4. SINGLE MARKET IDENTIFICATION WITH MULTIPLE ATTRIBUTES

**4.1. Marginal utility linear in unobserved taste.** We now turn to the main objective of the paper, which is to derive identifying shape restrictions in the multi attribute case of quality  $z \in \mathbb{R}^{d_z}$  and unobserved taste  $\varepsilon \in \mathbb{R}^{d_\varepsilon}$ , with  $d_z = d_\varepsilon > 1$ . We start with the case, where marginal utility is linear in unobserved taste.

**Assumption S2.** The surplus function is  $\zeta(x, \varepsilon, z) = z'\varepsilon$ .

Under Assumption S2, the consumer maximization problem is that of finding

$$V^*(x, \varepsilon) = \sup_z \{z'\varepsilon - V(x, z)\}. \quad (2)$$

Under suitable regularity, the first order condition yields:

$$\nabla_z V(x, z) = \varepsilon \quad (3)$$

and the demand function  $\varepsilon \mapsto V^*(x, \varepsilon)$  is, by definition, the convex conjugate (also known as Legendre-Fenchel transform) of  $V(x, z)$ . According to convex duality theory, the conjugate of  $V^*(x, \varepsilon)$  is  $V(x, z)$  itself if and only if  $V(x, z)$  is convex. Convexity of  $V$  will turn out to be the shape restriction that delivers identification in this case, where marginal utility is linear in unobserved taste.

**Assumption C2.** [Convexity restriction] The function  $V(x, z)$  is convex in  $z$  for all  $x$ .

Convexity of  $V$  in the univariate case of the previous section is equivalent to monotonicity of demand, or assortative matching, delivered by the single crossing shape restriction on  $\zeta$ . Since the function  $V(x, z)$  is equal to  $p(z) - \bar{U}(x, z)$ , it involves the endogenous price function. Hence, a condition on  $V$  is undesirable, since it constrains

non primitive quantities in the model. However, we show that, as monotonicity of inverse demand in the scalar case was implied by the single crossing condition, here, convexity of  $V$  is “essentially” always true under the maintained primitive conditions, in the sense of the following lemma.

**Lemma 1** (Convexity). *Under Assumptions EC, H, R and S2, for all  $x$ ,  $V(x, z) = (V^*)^*(x, z)$ ,  $P_{z|x}$  almost surely, where  $*$  denotes convex conjugation with respect to the second variable and hence  $(V^*)^*(x, z)$  is the convex envelope of  $V$  with respect to the second variable.*

The lemma implies in particular, that on any open subset of the support of  $P_{z|x}$ ,  $V(x, z)$  is a convex function of  $z$  as required.

As discussed in Section 3, positive assortative matching (monotonicity of demand) is difficult to extend to the multi-attribute case, but not the efficiency result that comes with positive assortative matching. Imagine a social planner maximizing total surplus over the distribution of heterogeneous consumers. The planner’s problem is to maximize  $\mathbb{E}[z'\varepsilon|x]$  over all possible allocations of qualities  $z$  to consumer types  $\varepsilon$ , i.e., over all pairs of random vectors  $(z, \varepsilon)$  under the constraint that the marginal distributions  $P_\varepsilon$  and  $P_{z|x}$  are fixed. One of the central results of optimal transport theory, Brenier’s Theorem (Theorem 1 in the Appendix), shows precisely that such a planner’s problem admits a unique pure allocation as solution, which takes the form of the inverse demand function  $\varepsilon = \nabla_z V(x, z)$  with  $V(x, z)$  convex in  $z$ . We see thereby that convexity of  $V(x, z)$  in  $z$  is the shape restriction that delivers identification as summarized in the following theorem, formally proved in the appendix.

**Theorem 1** (Identification with linear marginal utility in taste). *Under Assumptions EC, H, R, S2 and C2, the following statements hold:*

- (1)  $\bar{U}(x, z)$  is nonparametrically identified.

- (2) For all  $x \in X$ ,  $\bar{U}(x, z) = p(z) - V(x, z)$  and  $z \mapsto V(x, z)$  is the convex solution to the optimization problem  $\min_V (\mathbb{E}[V(x, z)|x] + \mathbb{E}[V^\zeta(x, \varepsilon)|x])$ , where  $\varepsilon \mapsto V^*(x, \varepsilon)$  is the convex conjugate of  $z \mapsto V(x, z)$ .

The structure of this identification proof is as follows (see the proof of Theorem 3, of which this is a special case). We show that there exists a unique allocation of qualities to tastes  $z \mapsto \varepsilon(x, z)$  that maximizes the consumer problem (2). A significant portion of the proof is dedicated to showing that  $z \mapsto V(x, z)$  is differentiable. Hence, once the allocation (inverse demand function)  $\varepsilon(x, z)$  is identified,  $V(x, z)$  satisfies the first order condition

$$\nabla_z V(x, z) = \nabla_z \zeta(x, \varepsilon(x, z), z), \quad P_{z|x}\text{-almost surely, for every } x.$$

The latter determines  $V(x, z)$ , and therefore  $\bar{U}(x, z)$ , up to a constant.

The identification is constructive, as  $V(x, z)$  can also be shown to be the convex solution to the minimization of  $\mathbb{E}[V(x, z)|x] + \mathbb{E}[V^*(x, \varepsilon)|x]$ , where  $V^*$  is the convex conjugate of  $V$ . This fact provides efficient computation strategies for inference on  $\bar{U}(x, z)$ , as shown for instance in Aurenhammer, Hoffmann, and Aronov (1998). The reasoning behind the identification result of Theorem 1 has implications beyond hedonic equilibrium models, as it provides identification conditions for a general nonlinear nonseparable simultaneous equations econometric model of the form  $z = f(x, \varepsilon)$ , where the vector of endogenous variables  $z$  has the same dimension as the vector of unobserved heterogeneity  $\varepsilon$ . Theorem 2 shows that in such models,  $f$  is nonparametrically identified within the class of gradients of convex functions.

**Theorem 2** (Nonlinear simultaneous equations). *In the simultaneous equations model  $z = f(x, \varepsilon)$ , with  $z, \varepsilon \in Z \subseteq \mathbb{R}^{d_z}$  and  $x \in X \subseteq \mathbb{R}^{d_x}$ , the function  $\varepsilon \mapsto f(x, \varepsilon)$  is identified under the following conditions.*

- (1)  $\varepsilon \mapsto f(x, \varepsilon)$  is the gradient of a convex function for all  $x \in X$ .

- (2) For all  $x \in X$ ,  $P_{\varepsilon|x}$  is fixed and is absolutely continuous with respect to Lebesgue measure and  $P_{\varepsilon}$  and  $P_{z|x}$  have finite variance.

In the univariate case, gradients of convex functions are the increasing functions, so that our identifying shape restriction directly generalizes monotonicity in Matzkin (2003).

**4.2. General case.** The identification result of Theorem 1 can be easily extended to allow for variation in the quality-unobserved taste interaction with observed type  $x$  as in  $\zeta(x, \varepsilon, z) = \phi(z)' \psi(x, \varepsilon)$ , where  $\phi$  and  $\psi$  are known functions and  $\phi$  is invertible. Going beyond this requires a new type of shape restriction that can be interpreted as a multivariate extension of the single crossing “Spence-Mirlees” condition.

Recalling our notation  $V(x, z) = p(z) - \bar{U}(x, z)$ , the consumer’s program is to choose quality vector  $z$  to maximize

$$\sup_z \{\zeta(x, \varepsilon, z) - V(x, z)\}. \quad (4)$$

In the one dimensional case, single crossing condition  $\zeta_{\varepsilon z}(x, \varepsilon, z) > 0$  delivered identification of inverse demand. We noted that the sign of the single crossing condition was not important for the identification result, rather the following implication of single crossing was.

**Assumption S3.** [Twist condition] For all  $x$  and  $z$ , the gradient  $\nabla_z \zeta(x, \varepsilon, z)$  of  $\zeta(x, \varepsilon, z)$  in  $z$  is injective as a function of  $\varepsilon$ .

Assumption S3, unlike the single crossing condition, is well defined in the multivariate case, and we shall show, using recent developments in optimal transport theory, that it continues to deliver the desired identification in the multivariate case. Before stating the theorem, we provide more intuition by further developing the parallel between this general case and the cases covered so far. Notice that the twist condition

of Assumption S3 is satisfied in the particular case of marginal utility linear in taste, as considered in Section 4.1 above.

Consider, as before, the hedonic market from the point of view of a social planner, who allocates qualities  $z$  to tastes  $\varepsilon$  in a way that maximizes total consumer surplus. The distribution of consumer tastes is  $P_\varepsilon$  and the distribution of qualities traded at equilibrium is  $P_{z|x}$ . For fixed observable type  $x$ , the variable surplus of a match between unobserved taste  $\varepsilon$  and quality  $z$  is  $\zeta(x, \varepsilon, z)$ . Hence, the planner's problem is to find an allocation of qualities to tastes, in the form of a joint probability  $\mu$  over the pair of random vectors  $(\varepsilon, z)$ , so as to maximize  $\mathbb{E}_\mu[\zeta(x, \varepsilon, z)|x]$  under the constraint that  $\varepsilon$  has marginal distribution  $P_\varepsilon$  and that  $z$  has marginal distribution  $P_{z|x}$ . This planner's problem

$$\max_{\mu} \mathbb{E}_\mu[\zeta(x, \varepsilon, z)|x] \quad \text{subject to } \varepsilon \sim P_\varepsilon, z \sim P_{z|x} \quad (5)$$

is equal to its dual

$$\min_{V, W} \mathbb{E}[W(x, \varepsilon)|x] + \mathbb{E}[V(x, z)|x] \quad \text{subject to } W(x, \varepsilon) + V(x, z) \geq \zeta(x, \varepsilon, z) \quad (6)$$

and both primal (5) and dual (6) are attained under the conditions of the Monge-Kantorovitch Theorem (Theorem 4 in the Appendix). Notice that the constraint in (6) can be written as

$$W(x, \varepsilon) = V^\zeta := \sup_z \{\zeta(x, \varepsilon, z) - V(x, z)\} \quad (7)$$

so that  $W(x, \varepsilon)$  is a candidate for the demand function mapping tastes  $\varepsilon$  into qualities  $z$  derived from the consumer's program (4). (7) defines a generalized notion of convex conjugation, which can be inverted, similarly to convex conjugation, into:

$$(V(x, z)^\zeta)^\zeta = \sup_\varepsilon \{\zeta(x, \varepsilon, z) - V^\zeta(x, \varepsilon)\}. \quad (8)$$

**Definition 2** ( $\zeta$ -convexity). *A function  $V$  is called  $\zeta$ -convex if and only if  $(V^\zeta)^\zeta = V$ .*



$\zeta$ -convexity, therefore, is a shape restriction that directly generalizes the convexity restriction of Assumption C2 in case of marginal utility linear in unobserved taste.

**Assumption C3.** [ $\zeta$ -convexity]  $V(x, z)$  is  $\zeta$ -convex as a function of  $z$  for all  $x$ .

As discussed in the previous section, Assumption C3 involves restrictions on the price function  $p(z)$ , which is endogenously determined at equilibrium. We show, however, that the twist condition Assumption S3 is the relevant shape restriction, corresponding to single crossing, and that  $\zeta$ -convexity follows, in the sense of the following lemma.

**Lemma 2.** *Under Assumptions EC, H, R and S3,  $V(x, z) = (V(x, z)^\zeta)^\zeta$ ,  $P_{z|x}$ -almost surely, for every  $x$ .*

Under Assumptions EC, H, R, S3 and C3, we show that there exists a unique allocation of qualities to tastes  $z \mapsto \varepsilon(x, z)$  that maximizes the consumer problem (7). Moreover, this allocation is such that markets clear, since  $\varepsilon(x, z)$  is distributed according to  $P_\varepsilon$  when  $z$  is distributed according to  $P_{z|x}$ . Heuristically, by the envelope theorem applied to  $V(x, z) = \sup_\varepsilon \{\zeta(x, \varepsilon, z) - V^\zeta(x, \varepsilon)\}$ , for a small variation  $\delta V^\zeta$  of  $V^\zeta$ , the variation in  $V$  is  $\delta V(x, z) = -\delta V^\zeta(x, \varepsilon(x, z), z)$ . Plugging the latter into the first order condition for (6) yields  $\mathbb{E}[\delta V^\zeta(x, \varepsilon)|x] = \mathbb{E}[\delta V^\zeta(x, \varepsilon(x, z))|x]$ . The latter holds for any small variation  $\delta V^\zeta$ , so that the distribution of allocation  $\varepsilon(x, z)$  is the same as the exogenously given distribution of unobserved tastes  $P_\varepsilon$ , so that the market clears.

Finally, once the allocation (inverse demand function)  $\varepsilon(x, z)$  is identified,  $V(x, z)$  satisfies the first order condition

$$\nabla_z V(x, z) = \nabla_z \zeta(x, \varepsilon(x, z), z), \quad P_{z|x}\text{-almost surely, for every } x.$$

The latter determines  $V(x, z)$ , and therefore  $\bar{U}(x, z)$ , up to a constant. We are now ready to state our main theorem, relating the Twist condition and the  $\zeta$ -convex shape restriction to nonparametric identification of preferences.

**Theorem 3** (Identification of preferences). *Under Assumptions EC, H, R, S3 and C3, the following statements hold:*

- (1)  $\bar{U}(x, z)$  is nonparametrically identified.
- (2) For all  $x \in X$ ,  $\bar{U}(x, z) = p(z) - V(x, z)$  and  $z \mapsto V(x, z)$  is the  $\zeta$ -convex solution to the optimization problem  $\min_V (\mathbb{E}[V(x, z)|x] + \mathbb{E}[V^\zeta(x, \varepsilon)|x])$ , where  $V^\zeta$  is defined in (7).

As before, the identification strategy is constructive and efficient computation of  $\bar{U}(x, z) = p(z) - V(x, z)$  is based on the identification of  $V$  as the solution to the optimization problem of Theorem 3(2). Again, the identification result of Theorem 3 has ramification beyond the framework of hedonic equilibrium models. Indeed, it provides the just identifying shape restriction for general consumer problems with multivariate unobserved preference heterogeneity, where consumers choose within a universe of goods, differentiated along more than one dimension. Theorem 3 tells us that the shape of interactions between good qualities and unobserved tastes governs the shape restriction that just identifies the utility function.

## 5. DISCUSSION

The identification results in the paper rely on observations from a single price schedule and the structural functions are just identified under normalization of the distribution of unobserved heterogeneity. Although  $\bar{U}(x, z)$  is not identified without such a normalization, or additional restrictions, there are features of preferences that are identified. Consider the model of Assumption S2.

$$U(x, \varepsilon, z) = \bar{U}(x, z) + z'\varepsilon.$$

From Theorem 6 in the Appendix, the inverse demand  $\varepsilon(x, z) = \nabla_z [p(z) - \bar{U}(x, z)]$  satisfies the following: for all bounded continuous functions  $\xi$ ,

$$\int \xi(\varepsilon) f_\varepsilon(\varepsilon) d\varepsilon = \int \xi (\nabla_z p(z) - \nabla_z \bar{U}(x, z)) f_{z|x}(z|x) dz. \quad (9)$$

Hence, taking  $\xi$  equal to the identity in (9) and assuming only that  $P_\varepsilon$  has mean zero, instead of fixing the whole distribution, yields identification of averaged partial effects  $\int \nabla_z \bar{U}(x, z) f_{z|x}(z|x) dz$  as

$$[\nabla_z \bar{U}(x, Z) | X = x] = \mathbb{E} [\nabla p(Z) | X = x]$$

from the fact that  $p(z)$  and  $f_{z|x}$  are identified.

Going beyond the latter result requires additional assumptions, which can take either of the following forms. (1) A separability assumption of the form  $\bar{U}(x, z) := z' \alpha(x) + \beta(z)$  can be imposed, in which case, the strategy outlined in Ekeland, Heckman, and Nesheim (2004) can be extended to the case of multiple attributes and yield identification without normalization of the distribution of unobserved tastes for quality  $\varepsilon$ . (2) Data from multiple markets can be brought to bear on the identification problem, or more generally a variable that shifts underlying distributions of producers and consumers, without affecting preferences and technology. To fix ideas, suppose two separate markets  $m_1$  and  $m_2$  (separate in the sense that producers, consumers or goods cannot move between markets) with underlying producer and consumer distributions  $(P_x^{m_1}, P_y^{m_1})$  and  $(P_x^{m_2}, P_y^{m_2})$ , are at equilibrium, with respective price schedules  $p^{m_1}(z)$  and  $p^{m_2}(z)$ . The additional identification assumption is that utility  $\bar{U}(x, z)$  and unobserved taste distribution  $P_\varepsilon$  are common across the two markets. In each market, we recover a nonparametrically identified utility function  $\bar{U}^m(x, z; P_\varepsilon)$ , where the dependence in the unknown distribution of tastes  $P_\varepsilon$  is emphasized. Therefore, the additional identifying restriction associated with multiple markets data takes the form  $\bar{U}^{m_1}(x, z; P_\varepsilon) = \bar{U}^{m_2}(x, z; P_\varepsilon)$  which defines an identified

set for the pair  $(\bar{U}, P_\varepsilon)$ . Further research is needed to characterize this identified set and derive conditions for point identification in this setting.

## 6. APPENDIX

Throughout the appendix, when there is no ambiguity, we drop the conditioning variable  $x$  from the notation and consider the theory of optimal transportation of distribution  $P_z$  of quality vector  $z \in \mathbb{R}^d$  to distribution  $P_\varepsilon$  of vector of unobserved tastes  $\varepsilon \in \mathbb{R}^d$ .

*Kantorovich problem.* We first consider the Kantorovich problem, which is the probabilistic allocation of qualities to tastes so as to maximize total surplus, where the surplus of a pair  $(\varepsilon, z)$  is given by the function  $\zeta(\varepsilon, z)$ , and the marginal distributions of qualities  $P_z$  and tastes  $P_\varepsilon$  are fixed constraints. We therefore define the set of allocations that satisfy the constraints as follows.

**Definition 3** (Probabilities with given marginals). *We denote  $\mathcal{M}(P_\varepsilon, P_z)$  the set of probability measures on  $\mathbb{R}^d \times \mathbb{R}^d$  with marginal distributions  $P_\varepsilon$  and  $P_z$ .*

With this definition, we can formally state the Kantorovich problem as follows.

$$(\text{PK}) = \sup_{\pi \in \mathcal{M}(P_\varepsilon, P_z)} \int \zeta(\varepsilon, z) d\pi(\varepsilon, z).$$

If we consider the special case of surpluses that are separable in  $\varepsilon$  and  $z$  and dominate  $\zeta(\varepsilon, z)$ , i.e., of the form  $W(\varepsilon) + V(z) \geq \zeta(\varepsilon, z)$ , the integral yields  $\int W(\varepsilon) dP_\varepsilon(\varepsilon) + \int V(z) dP_z(z)$ . We denote  $\Phi_\zeta$  the set of such functions.

**Definition 4** (Admissible set). *A pair of function  $(W, V)$  on  $\mathbb{R}^d$  belongs to the admissible set  $\Phi_\zeta$  if and only if  $W \in L^1(P_\varepsilon)$ ,  $V \in L^1(P_z)$  and  $W(\varepsilon) + V(z) \geq \zeta(\varepsilon, z)$  for  $P_\varepsilon$ -almost all  $\varepsilon$  and  $P_z$ -almost all  $z$ .*

The integral over separable surpluses

$$(\text{DK}) = \inf_{(W, V) \in \Phi_\zeta} \int W(\varepsilon) dP_\varepsilon(\varepsilon) + \int V(z) dP_z(z)$$

will in general yield a weakly larger total surplus than (PK), but it turns out that under very weak conditions, the two coincide.

**Theorem 4** (Kantorovich duality). *If  $\zeta$  is upper semi-continuous, then  $(\text{PK}) = (\text{DK})$  and there exists an allocation  $\pi \in \mathcal{M}(P_\varepsilon, P_z)$  that achieves the maximum in (PK).*

A proof of the Kantorovich duality theorem can be found in Chapter 1 of Villani (2003). We give here the intuition of the result based on switching infimum and supremum operations. First,  $\zeta$  is a continuous function, so that the mapping  $\mu \mapsto \int \zeta d\mu$  is weakly continuous. Since  $\mathcal{M}(P_\varepsilon, P_z)$  is weakly compact, the maximum in (PK) is achieved for some  $\pi$  by the Weierstrass Theorem. Hence, an optimal allocation exists. However, continuity of  $\zeta$  is not necessary.

To see the duality result, denote  $\chi_A(x) = 0$  if  $x \in A$  and  $-\infty$  otherwise. Then, we verify that

$$\chi_{\mathcal{M}(P_\varepsilon, P_z)} = \inf_{(W, V)} \left\{ \int W(z) dP_\varepsilon(\varepsilon) + \int V(z) dP_z(z) - \int (W(\varepsilon) + V(z)) d\pi(\varepsilon, z) \right\},$$

where the infimum is over all integrable functions, say. Now we can rewrite (PK) as follows:

$$(\text{PK}) = \sup_{\pi} \left\{ \int \zeta(\varepsilon, z) d\pi(\varepsilon, z) + \chi_{\mathcal{M}(P_\varepsilon, P_z)}(\pi) \right\},$$

where the supremum is taken over all joint probability measures. Assuming the infimum and supremum operations can be switched yields:

$$\begin{aligned} (\text{PK}) &= \inf_{(W, V)} \sup_{\pi} \left\{ \int \zeta(\varepsilon, z) d\pi(\varepsilon, z) \right. \\ &\quad \left. + \int W(z) dP_\varepsilon(\varepsilon) + \int V(z) dP_z(z) - \int (W(\varepsilon) + V(z)) d\pi(\varepsilon, z) \right\} \\ &= \inf_{(W, V)} \left\{ \int W(z) dP_\varepsilon(\varepsilon) + \int V(z) dP_z(z) \right. \\ &\quad \left. - \inf_{\pi} \int (W(\varepsilon) + V(z) - \zeta(\varepsilon, z)) d\pi(\varepsilon, z) \right\}. \end{aligned}$$

Consider the second infimum in the last display. If the function  $W(\varepsilon) + V(z) - \zeta(\varepsilon, z)$  takes a negative value, then, choosing for  $\pi$  the Dirac mass at that point will yield an infimum of  $-\infty$ . Therefore, we have:

$$\inf_{\pi} \int (W(\varepsilon) + V(z) - \zeta(\varepsilon, z)) d\pi(\varepsilon, z) = \chi_{\Phi_\zeta}(W, V),$$

so that

$$(\text{PK}) = \inf_{(W, V)} \left\{ \int W(z) dP_\varepsilon(\varepsilon) + \int V(z) dP_z(z) - \chi_{\Phi_\zeta}(W, V) \right\} = (\text{DK})$$

as required.

We now see that the dual is also achieved.

**Theorem 5** (Kantorovich duality (continued)). *If  $(PK) < \infty$  and there exist integrable functions  $\zeta^\varepsilon$  and  $\zeta^z$  such that  $\zeta(\varepsilon, z) \geq \zeta^\varepsilon(\varepsilon) + \zeta^z(z)$ , then there exists a  $\zeta$ -convex function (see Definition 2)  $V$*

such that

$$\int V^\zeta(\varepsilon)dP_\varepsilon(\varepsilon) + \int V(z)dP_z(z)$$

achieves (DK). In addition, if  $\pi$  is an optimal allocation, i.e., achieves (PK), and  $(V^\zeta, V)$  is an optimal  $\zeta$ -conjugate pair, i.e., achieves (DK), then

$$V^\zeta(\varepsilon) + V(z) \geq \zeta(\varepsilon, z) \text{ with equality } \pi\text{-almost surely.} \quad (10)$$

*Idea of the proof.* The proof can be found in Chapter 5 of Villani (2009). The last statement of Theorem 5 is easy to see. If  $\pi$  achieves (PK) and  $(V^\zeta, V)$  achieves (DK), then, as (PK)=(DK) by Theorem 4, we have  $\int [V^\zeta(\varepsilon) + V(z) - \zeta(\varepsilon, z)]d\pi(\varepsilon, z) = 0$ . The integrand is non negative, since  $(V^\zeta, V) \in \Phi_\zeta$ . Hence,  $V^\zeta + V = \zeta$ ,  $\pi$ -almost surely, as desired. The proof of existence of an optimal pair of  $\zeta$ -convex functions achieving (DK) revolves around the notion of cyclical monotonicity.

In view of the above, if  $(\phi, \psi)$  achieve (DK) and a sequence of pairs  $(\varepsilon_i, z_i)_{i=1, \dots, m}$  belong to the support of the optimal allocation  $\pi$ , then  $\phi(\varepsilon_i) + \psi(z_i) = \zeta(\varepsilon_i, z_i)$  for each  $i = 1, \dots, m$ . On the other hand, since  $(\phi, \psi) \in \Phi_\zeta$ , we have  $\phi(\varepsilon_i) + \psi(z_{i+1}) \geq \zeta(\varepsilon_i, z_{i+1})$  for each  $i = 1, \dots, m-1$ , and  $\phi(\varepsilon_m) + \psi(z) \geq \zeta(\varepsilon_m, z)$  for an arbitrary  $z$ . Subtracting and adding up yields  $\psi(z) \geq \psi(z_1) + [\zeta(\varepsilon_m, z) - \zeta(\varepsilon_m, z_m)] + \dots + [\zeta(\varepsilon_1, z_2) - \zeta(\varepsilon_1, z_1)]$ . Since the functions in the pair  $(\phi, \psi)$  are only determined up to a constant, normalize  $\psi(z_1) = 0$  and define  $V$  as the supremum of all functions  $\psi$  satisfying  $\psi(z) \geq [\zeta(\varepsilon_m, z) - \zeta(\varepsilon_m, z_m)] + \dots + [\zeta(\varepsilon_1, z_2) - \zeta(\varepsilon_1, z_1)]$  over all choices of  $(\varepsilon_i, z_i)_{i=1, \dots, m}$  in the support of  $\pi$  and all  $m \geq 0$ . It turns out that  $V^\zeta(\varepsilon) + V(z) = \zeta(\varepsilon, z)$ ,  $\pi$ -almost surely, so that integration over  $\pi$  yields the fact that  $(V^\zeta, V)$  achieves (DK) as desired.  $\square$

*The quadratic case and Brenier's Theorem.* In the special case of Assumption S2, where  $\zeta(\varepsilon, z) = z'\varepsilon$ , the planner's program (PK) writes

$$\sup_{\pi \in \mathcal{M}(P_\varepsilon, P_z)} \int z'\varepsilon d\pi(\varepsilon, z)$$

and the set  $\Phi_\zeta$  becomes

$$\Phi = \{(W, V) : W(\varepsilon) + V(z) \geq z'\varepsilon\}.$$

The pair  $(W, W^*) \in \Phi$  defined by

$$\begin{aligned} W^*(z) &= \sup_{\varepsilon} \{z'\varepsilon - W(\varepsilon)\}, \\ W(\varepsilon) &= \sup_z \{z'\varepsilon - W^*(z)\} \end{aligned}$$

achieves the minimum in the dual problem (DK). Notice that  $W$  and  $W^*$  are standard Fenchel-Legendre convex conjugates of each other and that  $W = W^{**}$  and is hence convex.

In this case,  $\nabla W^*(z) = \nabla_z \zeta(\varepsilon, z)$  simplifies to  $\nabla W^*(z) = \varepsilon$ , which guarantees uniqueness and purity of the optimal assignment  $z = \nabla W(\varepsilon)$ , where  $V$  is convex. As a corollary,  $\nabla W$  is a  $P_\varepsilon$ -almost surely uniquely determined gradient of a convex function.

**Theorem 6** (Brenier). *Suppose  $P_\varepsilon$  is absolutely continuous with respect to Lebesgue measure and that  $P_\varepsilon$  and  $P_z$  have finite second order moments. Then, there exists a  $P_\varepsilon$ -almost surely unique map of the form  $\nabla W$ , where  $W$  is convex, such that  $\int \varepsilon' \nabla W(\varepsilon) dP_\varepsilon(\varepsilon)$  achieves the maximum in (PK) with  $\zeta(\varepsilon, z) = z' \varepsilon$ .  $W$  is the  $P_\varepsilon$ -almost everywhere uniquely determined convex map such that  $P_\varepsilon(\nabla W^{-1}(B)) = P_z(B)$  for all Borel subsets  $B$  of the support of  $P_z$ . Moreover,  $(W, W^*)$  achieves the dual program (DK).*

A proof of Theorem 6 can be found in Section 2.1.5 of Villani (2003). Note that  $W$  is not only the unique convex map that solves the optimization problem.  $\nabla W$  is the unique gradient of a convex map that pushes forward probability measure  $P_\varepsilon$  to  $P_z$ . Hence, identification can be achieved for a nonlinear simultaneous equations model  $z = f(x, \varepsilon)$  without an underlying assumption about how choices  $z$  were generated from tastes  $(x, \varepsilon)$ . This is the content of Theorem 2, which is therefore seen to be a straightforward application of Theorem 6.

*Proof of Lemma 1.* This is a corollary of Lemma 2, since Assumptions S2 and C2 imply Assumptions S3 and C3. □

*Proof of Theorem 1.* This is a corollary of Theorem 3, since Assumptions S2 and C2 imply Assumptions S3 and C3. □

*Proof of Theorem 2.* Since  $z = f(x, \varepsilon)$  and  $X \perp\!\!\!\perp \varepsilon$ , for any Borel set  $B$ ,

$$\begin{aligned} \mathbb{P}(Z \in B | X = x) &= \mathbb{P}(f(x, \varepsilon) \in B | X = x) \\ &= P_\varepsilon(f(x, \varepsilon) \in B). \end{aligned} \tag{11}$$

By Theorem 6, a gradient of a convex function satisfying (11) is  $P_\varepsilon$ -almost everywhere uniquely determined, hence the result. □

*Proof of Lemma 2.* By definition of  $V^\zeta$ , we have

$$V(x, z) \geq \zeta(x, \varepsilon, z) - V^\zeta(x, \varepsilon) \quad (12)$$

By Assumption EC, there is equality in (12)  $\gamma$ -almost everywhere, where  $\gamma$  is an hedonic equilibrium measure on  $\tilde{X} \times Z \times \tilde{Y}$ . As, by definition of  $\zeta$ -conjugation,

$$V^{\zeta\zeta}(x, z) = \sup_{\varepsilon} [\zeta(x, \varepsilon, z) - V^\zeta(x, \varepsilon)]$$

we have

$$V(x, z) \geq V^{\zeta\zeta}(x, z), \quad (13)$$

by taking supremum over  $\varepsilon$  in (12).

Now, for  $P_{z|x}$  almost every  $z$ , there is an  $\varepsilon$  such that  $(z, \varepsilon)$  is in the support of  $\gamma$ , and so, for this  $\varepsilon$ , we have equality in (12)

$$V(x, z) = \zeta(x, \varepsilon, z) - V^\zeta(x, \varepsilon).$$

But the right hand side is bounded above by  $V^{\zeta\zeta}(x, z)$  by definition, so we get

$$V(x, z) \leq V^{\zeta\zeta}(x, z).$$

Combined with (13), this tells us  $V(x, z) = V^{\zeta\zeta}(x, z)$ ,  $P_{z|x}$  almost everywhere.  $\square$

*Proof of Theorem 3(1).* Step 1: identification of inverse demand. For a fixed observable type  $x$ , assume that the types  $(x, \varepsilon_0)$  and  $(x, \varepsilon_1)$  both choose the same good,  $\bar{z} \in Z$ , from producers  $\tilde{y}_0$  and  $\tilde{y}_1$ , respectively.

We want to prove that this implies the unobservable types are also the same; that is, that  $\varepsilon_0 = \varepsilon_1$ . This property is equivalent to having a map from the goods  $Z$  to the unobservable types, for each fixed observable type.

Note that  $\bar{z}$  must maximize the joint surplus for both  $\varepsilon_0$  and  $\varepsilon_1$ . That is, setting

$$S(x, \varepsilon, \tilde{y}) = \max_z [\bar{U}(x, z) + \zeta(x, \varepsilon, z) - C(\tilde{y}, z)] \quad (14)$$

we have,

$$S(x, \varepsilon_0, \tilde{y}_0) = \bar{U}(x, \bar{z}) + \zeta(x, \varepsilon_0, \bar{z}) - C(\tilde{y}_0, \bar{z}) \quad (15)$$

and

$$S(x, \varepsilon_1, \tilde{y}_1) = \bar{U}(x, \bar{z}) + \zeta(x, \varepsilon_1, \bar{z}) - C(\tilde{y}_1, \bar{z}). \quad (16)$$



By Assumption EC, we can apply Lemma 1 of Chiappori, McCann, and Nesheim (2010), so that the pair of indirect utilities  $(V, W)$ , where

$$\begin{aligned} V(\tilde{x}) &= \sup_{z \in Z} (U(\tilde{x}, z) - p(z)) \\ W(\tilde{y}) &= \sup_{z \in Z} (p(z) - C(\tilde{y}, z)), \end{aligned}$$

achieve the dual (DK) of the optimal transportation problem

$$\sup_{\pi \in \mathcal{M}(P_{\tilde{x}}, P_{\tilde{y}})} \int S(\tilde{x}, \tilde{y}) d\pi(\tilde{x}, \tilde{y}),$$

with solution  $\pi$ . This implies, from Theorem 5, that for  $\pi$ -almost any pairs  $(\tilde{x}_0, \tilde{y}_0)$  and  $(\tilde{x}_1, \tilde{y}_1)$ ,

$$\begin{aligned} V(\tilde{x}_0) + W(\tilde{y}_0) &= S(\tilde{x}_0, \tilde{y}_0), \\ V(\tilde{x}_1) + W(\tilde{y}_1) &= S(\tilde{x}_1, \tilde{y}_1), \\ V(\tilde{x}_0) + W(\tilde{y}_1) &\geq S(\tilde{x}_0, \tilde{y}_1), \\ V(\tilde{x}_1) + W(\tilde{y}_0) &\geq S(\tilde{x}_1, \tilde{y}_0). \end{aligned}$$

We therefore deduce the condition (called the 2-monotonicity condition):

$$S(x, \varepsilon_0, \tilde{y}_0) + S(x, \varepsilon_1, \tilde{y}_1) \geq S(x, \varepsilon_1, \tilde{y}_0) + S(x, \varepsilon_0, \tilde{y}_1),$$

recalling that  $\tilde{x} = (x, \varepsilon)$ . Now, by definition of  $S$  as the maximized surplus, we have

$$S(x, \varepsilon_1, \tilde{y}_0) \geq \bar{U}(x, \bar{z}) + \zeta(x, \varepsilon_1, \bar{z}) - C(\tilde{y}_0, \bar{z}) \quad (17)$$

and

$$S(x, \varepsilon_0, \tilde{y}_1) \geq \bar{U}(x, \bar{z}) + \zeta(x, \varepsilon_0, \bar{z}) - C(\tilde{y}_1, \bar{z}). \quad (18)$$

Inserting this, as well as (15) and (16) into the 2-monotonicity inequality yields

$$\begin{aligned} \bar{U}(x, \bar{z}) + \zeta(x, \varepsilon_0, \bar{z}) - C(\tilde{y}_0, \bar{z}) &+ \bar{U}(x, \bar{z}) + \zeta(x, \varepsilon_1, \bar{z}) - C(\tilde{y}_1, \bar{z}) \\ &\geq S(x, \varepsilon_1, \tilde{y}_0) + S(x, \varepsilon_0, \tilde{y}_1) \\ &\geq \bar{U}(x, \bar{z}) + \zeta(x, \varepsilon_1, \bar{z}) - C(\tilde{y}_0, \bar{z}) + \bar{U}(x, \bar{z}) + \zeta(x, \varepsilon_0, \bar{z}) - C(\tilde{y}_1, \bar{z}). \end{aligned}$$

But the left and right hand sides of the preceding string of inequalities are identical, so we must have equality throughout. In particular, we must have equality in (17) and (18). Equality in (17), for example, means that  $\bar{z}$  maximizes  $z \mapsto \bar{U}(x, z) + \zeta(x, \varepsilon_1, z) - C(\tilde{y}_0, z)$ , and so, as  $\bar{z}$  is in the interior of  $Z$  by Assumption EC, we have

$$\nabla_z \zeta(x, \varepsilon_1, \bar{z}) = \nabla_z C(\tilde{y}_0, \bar{z}) - \nabla_z \bar{U}(x, \bar{z}). \quad (19)$$

But now recall :  $\bar{z}$  also maximizes  $z \mapsto \bar{U}(x, z) + \zeta(x, \varepsilon_0, z) - C(\tilde{y}_0, z)$ , and so also have

$$\nabla_z \zeta(x, \varepsilon_0, \bar{z}) = \nabla_z C(\tilde{y}_0, \bar{z}) - \nabla_z \bar{U}(x, \bar{z}). \quad (20)$$

Equations (19) and (20) then imply

$$\nabla_z \zeta(x, \varepsilon_1, \bar{z}) = \nabla_z \zeta(x, \varepsilon_0, \bar{z})$$

and Assumption S3 implies  $\varepsilon_1 = \varepsilon_0$ .

Step 2: Differentiability of  $V$  in  $z$ . The method of proof of Step 2 is to prove that the subdifferential at each  $z_0$  is a singleton, which is equivalent to differentiability at  $z_0$ .

**Definition 5** (Subdifferential). *The subdifferential  $\partial\psi(x_0)$  of a function  $\psi : \mathbb{R}^d \rightarrow \mathbb{R} \cup \{+\infty\}$  at  $x_0 \in \mathbb{R}^d$  is the set of vectors  $p \in \mathbb{R}^d$ , called subgradients, such that  $\psi(x) - \psi(x_0) \geq p'(x - x_0) + o(|x - x_0|)$ .*

From Assumption C3,  $V(x, z)$  is  $\zeta$ -convex, and hence locally semiconvex, by Proposition C.2 in Gangbo and McCann (1996). We recall the definition of local semiconvexity from the latter paper.

**Definition 6** (Local semiconvexity). *A function  $\psi : \mathbb{R}^d \rightarrow \mathbb{R} \cup \{+\infty\}$  is called semiconvex at  $x_0 \in \mathbb{R}^d$  if there is a scalar  $\lambda > 0$  such that  $\psi(x) + \lambda|x|^2$  is convex on some open ball centered at  $x_0$ .*

Since the term  $\lambda|x|^2$  in the definition of local semiconvexity simply shifts the subdifferential by  $2\lambda x$ , we can extend Theorem 25.6 in Rockafellar (1970) to locally semiconvex functions and obtain the following lemma.

**Lemma 3.** *Let  $\psi : \mathbb{R}^d \rightarrow \mathbb{R} \cup \{+\infty\}$  be a locally semiconvex function, and suppose that  $q \in \mathbb{R}^d$  is an extremal point in the subdifferential  $\partial\psi(x_0)$  of  $\psi$  at  $x_0$ . Then there exist a sequence  $x_n$  converging to  $x_0$ , such that  $\psi$  is differentiable at each  $x_n$  and the gradient  $\nabla\psi(x_n)$  converges to  $q$ .*

Now, Step 1 shows that for each fixed  $z$ , the set

$$\{\varepsilon : V(x, z) + V^\zeta(x, \varepsilon) = \zeta(x, \varepsilon, z)\} := \{f(z)\} \quad (21)$$

is a singleton. We claim that this means  $V$  is differentiable everywhere. Fix a point  $z_0$ . We will prove that the subdifferential  $\partial V(z_0)$  contains only one extremal point, which will imply the desired result (as this means the subdifferential must be a singleton, and so  $V$  must be differentiable at  $z_0$ ).

Let  $q$  be any extremal point in  $\partial V(z_0)$ . Let  $z_n$  be a sequence satisfying the conclusion in Lemma 3. Now, as  $V$  is differentiable at each point  $z_n$ , we have the envelope condition

$$\nabla V_z(x, z_n) = \nabla_z \zeta(x, \varepsilon_n, z_n) \quad (22)$$

where  $\varepsilon_n = f(z_n)$  is the unique point giving equality in (21).

By Assumption R(3), the inverse mapping  $[\nabla_z \zeta(x, \cdot, z)]^{-1}$  is locally Lipschitz, and since the  $\nabla_z V(x, z_n)$  converge, the points  $\varepsilon_n = [\nabla_z \zeta(x, \cdot, z)]^{-1}(\nabla_z V(x, z_n))$  remain in a bounded set. We can therefore pass to a convergent subsequence  $\varepsilon_n \rightarrow \varepsilon_0$ .

By continuity of  $\nabla_z \zeta$ , we can pass to the limit in (22) and, recalling that the left hand side tends to  $q$ , we obtain

$$q = \nabla_z \zeta(x, \varepsilon_0, z_0).$$

Now, as we have the equality

$$V(x, z_n) + V^\zeta(x, \varepsilon_n) = \zeta(x, \varepsilon_n, z_n)$$

we can pass to the limit to obtain

$$V(x, z_0) + V^\zeta(x, \varepsilon_0) = \zeta(x, \varepsilon_0, z_0).$$

But this means  $\varepsilon_0 = f(z_0)$ , and so  $q = \nabla_z \zeta(x, \varepsilon_0, z_0) = \nabla_z \zeta(x, f(z_0), z_0)$  is uniquely determined by  $z_0$ . This means that the subdifferential can only have one extremal point, completing the proof of differentiability of  $V$ .

Step 3: Since by Step 2,  $V(x, z) = p(z) - \bar{U}(x, z)$  is differentiable, and by Step 1, the inverse demand function  $\varepsilon(x, z)$  is uniquely determined, the first order condition  $\nabla_z \zeta(x, \varepsilon(x, z), z) = \nabla p(z) - \nabla_z \bar{U}(x, z)$  identifies  $\nabla_z \bar{U}(x, z)$  as required.  $\square$

*Proof of Theorem 3(2).* In Part (1), we have shown uniqueness (up to location) of the pair  $(V, V^\zeta)$  such that  $V(x, z) + V^\zeta(x, \varepsilon) = \zeta(x, \varepsilon, z)$ ,  $\pi$ -almost surely. By Theorem 5, this implies that  $(V, V^\zeta)$  is the unique (up to location) pair of  $\zeta$ -conjugates that solves the dual Kantorovitch problem as required.

## REFERENCES

- AURENHAMMER, F., F. HOFFMANN, and B. ARONOV (1998): “Minkowski-type theorems and least squares clustering,” *Algorithmica*, 20, 61–76.
- BAJARI, P., and L. BENKARD (2005): “Demand estimation with heterogeneous consumers and unobserved product characteristics: a hedonic approach,” *Journal of Political Economy*, 113, 1239–1276.

- BARTIK, T. (1987): “The estimation of demand parameters in hedonic price models,” *Journal of Political Economy*, 95, 81–88.
- BECKER, G. (1973): “A Theory of marriage: part I,” *Journal of Political Economy*, 81, 813–846.
- BISHOP, K., and C. TIMMINS (2011): “Hedonic prices and implicit markets: Estimating marginal willingness to pay for differentiated products without instrumental variables,” *NBER Working Paper Series*, 17611, 1–38.
- BROWN, J., and H. ROSEN (1982): “On the estimation of structural hedonic price models,” *Econometrica*, 50, 765–769.
- CHIAPPORI, P.-A., R. MCCANN, and L. NESHEIM (2010): “Hedonic price equilibria, stable matching, and optimal transport: equivalence, topology, and uniqueness,” *Economic Theory*, 42, 317–354.
- CHIONG, K., A. GALICHON, and M. SHUM (2013): “Estimating dynamic discrete choice models via convex analysis,” unpublished manuscript.
- CHOO, E., and A. SIOW (2006): “Who marries whom and why?,” *Journal of Political Economy*, 114, 175–201.
- DUPUY, A., A. GALICHON, and M. HENRY (2014): “Entropy methods for identifying hedonic models,” *Mathematics and Financial Economics*, 8(4), 405–416.
- EKELAND, I. (2010): “Existence, uniqueness and efficiency of equilibrium in hedonic markets with multidimensional types,” *Economic Theory*, 42, 275–315.
- EKELAND, I., J. HECKMAN, and L. NESHEIM (2004): “Identification and estimation of hedonic models,” *Journal of Political Economy*, 112, 60–109.
- EPPLE, D. (1987): “Hedonic prices and implicit markets: Estimating demand and supply functions for differentiated products,” *Journal of Political Economy*, 95, 59–80.
- GALICHON, A., and B. SALANIÉ (2012): “Cupid’s invisible hand; social surplus and identification in matching models,” unpublished manuscript.
- GANGBO, W., and R. MCCANN (1996): “The geometry of optimal transportation,” *Acta Mathematica*, 177, 113–161.
- HARDY, G., J. LITTLEWOOD, and G. PÓLYA (1952): *Inequalities*. Cambridge University Press.
- HECKMAN, J., R. MATZKIN, and L. NESHEIM (2010): “Nonparametric identification and estimation of nonadditive hedonic models,” *Econometrica*, 78(5), 1569–1591.
- KHAN, S., and K. LANG (1988): “Efficient estimation of structural hedonic systems,” *International Economic Review*, 29, 157–166.

- KNOTHE, H. (1957): “Contributions to the theory of convex bodies,” *Michigan Mathematical Journal*, 4, 39–52.
- MATZKIN, R. (2003): “Nonparametric estimation of nonadditive random functions,” *Econometrica*, 71, 1339–1375.
- MATZKIN, R. (2013): “Nonparametric identification in structural econometric models,” *Annual Review of Economics*, 5, 457–486.
- NESHEIM, L. (2013): “Identification of multidimensional hedonic models,” Conference Presentation.
- ROCKAFELLAR, T. (1970): *Convex Analysis*. Princeton University Press, Princeton.
- ROSEN, S. (1974): “Hedonic prices and implicit markets: product differentiation in pure competition,” *Journal of Political Economy*, 82, 34–55.
- ROSENBLATT, M. (1952): “Remarks on a multivariate transformation,” *The Annals of Mathematical Statistics*, 23, 470–472.
- TINBERGEN, J. (1956): “On the theory of income distribution,” *Weltwirtschaftliches Archiv*, 77, 155–173.
- VILLANI, C. (2003): *Topics in Optimal Transportation*. American Mathematical Society, Providence.
- VILLANI, C. (2009): *Optimal Transport*. Springer, Berlin.