Contents

1

Introduction: Matching Models in Economics

Matching models can be (and have been) applied to a host of contexts, from finance to the labor market. In this book, however, I shall mostly illustrate my arguments with examples coming from family economics. And I shall start with two related puzzles that are directly related to family formation (and dissolution) and to intrafamily allocation, although their scope is obviously much larger. One has to do with the increase in inequality in the US in recent decades, and the other with some remarkable trends in gender-specific demand for higher education.

1.1 Motivation: Two Puzzles

1.1.1 Inequality

Rising inequality has been one of the most striking features of the recent decades. In the US, according to Kenworthy and Smeeding (2013), the Gini coefficients of both pre- and after-tax income distributions grew by 10 percentage points between 1979 and 2007 (from 0.39 to 0.49 and from 0.35 to 0.45, respectively). Similarly, the pre-tax income share of the top 10% of the US income distribution, which remained below 35% during the three decades following World War II, sharply increased from the early 1980s on, reaching 50% in 2010 (Saez 2013). Rising inequality is a complex phenomenon that has several causes, from skill-biased technological change to globalization. But one of these causes—although probably not the dominant one—is directly linked to family economics. Indeed, several papers have argued that some long-term demographic trends have affected the evolution of inequality; chief among these trends are the rise of single-parent families and an increase in assortative matching for existing couples.[1]

These demographic effects, whose exact magnitude is still discussed, unavoidably generate some difficult questions about the concepts at stake. An increase in assortative matching is not easy to define. We must first define the trait we should concentrate on: are we talking about education? wages? income? Income is probably a disputable choice, if only because it is obviously endogenous—the same matching patterns may exhibit strikingly different assortativeness in income if

[1] For instance, according to Greenwood et al. (2014): "if people matched in 2005 according to the 1960 standardized mating pattern there would be a significant reduction in income inequality; i.e., the Gini drops from 0.43 to 0.35" (p. 352).

some exogenous shock affects spouses' labor supply and participation in the labor market. Human capital is a better choice, since it can reasonably be considered as (at least partly) given when people enter the marriage market; note, however, that the impact of human capital investments on matching prospects may play a central role in agents' decisions regarding education (more on this later).

Second, while any increase in the percentage of single people has a mechanical impact on our *measures* of inequality, the economic reality behind the estimates is harder to assess. Consider a couple with annual income equal to $30,000—well above the poverty level.[2] Assume, now, that the couple divorces; we have two singles with respective incomes $12,000 and $18,000—one of which is below the poverty threshold. According to standard measures, both poverty and inequality have increased. Still, and even if we ignore the role of public consumptions and economies of scale, the actual *change* in economic inequality must unavoidably depend on the inequality that prevailed *within* the pre-divorce couple. Standard inequality measures systematically ignore this inequality; in order to compare income across families of different composition, they use equivalence scales, which essentially amounts to assuming fair intracouple distribution.[3] That, however, does not mean that intrafamily inequality does not exist; in fact, it has been argued that such issues should play a central role in both the measure of inequality and the definition of policies aiming at reducing it.[4] An immediate difficulty, however, is that intrafamily inequality is notoriously hard to measure. As we shall see, matching models may in many cases provide a very useful tool for this purpose.

Third, the mere notion of "increase in assortativeness" is somewhat vague. In recent decades, women's education has changed drastically; for instance, the percentage of women aged 30 to 40 with a postgraduate degree has surged from 3% to more than 11% over the last 40 years. A mechanical effect of this increase is that highly educated men are more likely now to marry a highly educated wife. But are the observed changes in matching patterns only due to the mechanical effects of variations in the distributions of male and female education? Or did preferences for assortativeness also increase over the period? More fundamentally, how can such "preferences for assortativeness" be defined, and how can we measure their variations?

1.1.2 Demand for Higher Education

As I just mentioned, one of the salient trends in recent decades is the increased investment in education by women and the closing (actually reversing) of the gap

---

[2] According to the US Census Bureau, in 2014 the poverty threshold was equal to $15,853 for a childless couple and $12,316 for a childless single.

[3] See, for instance, Chiappori (2016).

[4] See, for instance, Chiappori and Meghir (2015).

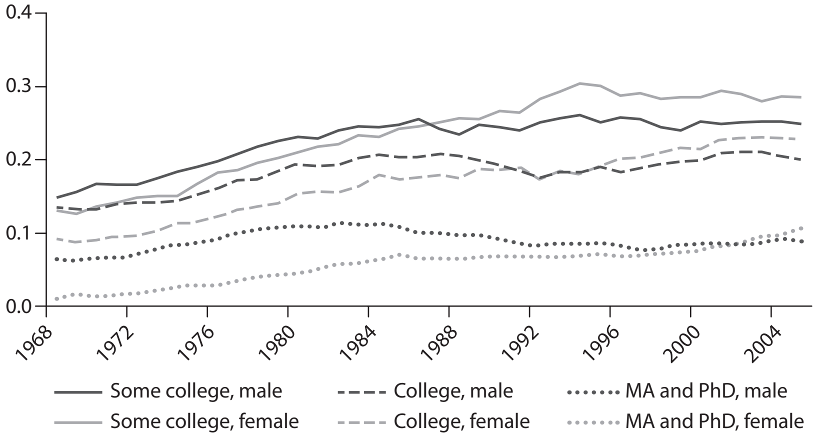

Figure 1.1. Educational attainment in US of spouses by husbands' year of birth (Chiappori, Iyigun, and Weiss 2009).

in schooling between men and women. Goldin and Katz (2008) show that, starting with the 1970 birth cohort, women have attained higher college graduation rates than men in the United States. They find similar reversals in 15 OECD countries.

Figure 1.1, borrowed from Chiappori, Iyigun, and Weiss (2009), describes the time trends in levels of school completion for men and women, aged 30 to 40, for the United States. As can be seen, the proportions of women with some college education, college completion, and advanced degrees (M.A., Ph.D.) have increased much faster than the corresponding proportions for men. By 2003, women had overtaken men in all of these three categories.

A particularly interesting feature pointed out by this graph is the asymmetry between genders regarding demand for higher education over the last few decades. Consider the percentage of individuals with a postgraduate degree, in dotted lines. That the labor market reward for graduate education has been increasing over the last four decades (and particularly at the end of the twentieth century) has been abundantly documented. When the return on some investment (here in human capital) sharply increases, one would expect the demand for such investment to surge. Indeed, we see that the percentage of women with a postgraduate degree more than trebles over the period—from about 3% in the mid-1970s to more than 11% in the mid-2000s. But no such increase is visible for men; if anything, the corresponding proportion slightly *declines* over the period.

These features raise two broad questions. First, how can we explain this puzzling asymmetry? In other words, how is it that men and women, faced with the same incentives, react in totally opposite ways? Second, and probably more importantly, what will be the long-term consequences of these evolutions? Changes in female education, labor supply, and earnings have been nothing short of dramatic over the period in the US but also in almost all developed countries, and actually in many intermediate or developing economies as well. It is hard to believe that such a landslide will not have deep and lasting effects on the distribution of powers within households, therefore ultimately on household behavior. This impact may be particularly important in developing countries, where we have considerable evidence that intrahousehold distribution of decision power strongly affects actual choices, including for such crucial issues as investment in children's human capital (health, education, etc.).

The two examples just described—and many others—share a common feature: they cannot be discussed without referring to an explicit model of household behavior formation, dissolution, and decision making. We need to understand the determinants of assortative matching (which factors make people more likely to marry their own kind, and how do these factors—or their impact—vary across time?), household formation and separation (why did the number of singles increase so drastically over the period?), and intrafamily allocation of resources and decision power. The purpose of this book is precisely to provide a framework in which these questions can be analyzed from both a theoretical and an empirical perspective.

1.2 Matching Models: Main Features

Matching models exist in various versions. All of these present some common characteristics, but they also differ in many respects. While a detailed presentation of the various models will be offered in the next chapter, it is useful to briefly recall their main features.

1.2.1 Absence of Frictions

First of all, matching models assume a frictionless environment. This means that, in the matching process between, say, firms and CEOs, each firm is assumed to have free access to the pool of all potential CEOs, with perfect knowledge of the characteristics of each of them, and conversely. In other words, matching models disregard the cost of acquiring information about potential matches, as well as the role of meeting technologies of all sorts (from social media to head hunters and from clubs to pure luck).

As always, an assumption of this kind is but a simplification of the infinite complexity of real-life processes. The question is whether such a simplification is acceptable. While the answer can only be case-specific, two general remarks can be made. First, the relevance of a frictionless setting largely depends on the question under consideration. A labor economist who is mostly interested in the dynamics of unemployment may be rightly reluctant to use a frictionless framework for modeling the matching of workers to jobs: there is a general consensus that a large fraction of unemployment results from various frictions on the labor market, and a search model should probably be preferred. If, however, the main issue is the allocation of specific individuals to specific tasks—say, which type of CEO ends up heading which type of firms—then a frictionless context may be more acceptable: while it is probably true that firms may not have a perfect knowledge of the pool of available CEOs, the resulting bias may be of second order, and neglecting it may be fully appropriate.

Second, the size of the market, as well as its structure, may have an impact on the relevance of a particular framework. Matching within a small community, in which individuals know each other, may be closer to the frictionless reference than in a very large market. Even with large populations (say, the "market for marriage" in a region of the US), one may want to concentrate on a very specific aspect involving a small number of broad categories (say, matching by education, broadly defined by four or five possible levels); then one may, for the sake of simplicity, disregard other individual traits and consider that two spouses with the same education are perfect substitutes. In such a context, neglecting frictions may be admissible.[5]

In the end, the choice of a specific model (frictionless matching versus search or other) should be driven, at least in part, by empirical considerations: what are the main stylized features of the situation we want to investigate, what types of frictions are more likely to matter, what importance should they represent, and last but not least which version fits the data best. It is therefore crucial to (i) deeply understand the meaning and implications of the various models under consideration, including their auxiliary hypothesis and apparently mundane details; and (ii) keep in mind the issues related to the empirical implementation of the various concepts at stake.

1.2.2 Are Transfers Relevant?

Within the family of frictionless matching frameworks, a second and crucial distinction relies on the role of transfers. The issue, here, is whether a technology exists that would allow one to transfer utility between agents participating to a matching process. Whether such transfers are possible makes a fundamental difference, because, when available, they allow agents to *bid* for their preferred mate by accepting the reduction of own gain from the match in order to increase the partner's. The exact nature of these bids depends on the context and may not take

[5] One could add that a small but important literature considers the issue of convergence of search equilibria to stable matchings of the frictionless framework when search frictions become negligible. The issue is quite complex, as one can find examples where the limit of search equilibria fails to be efficient (see, for instance, Lauermann and Noldeke 2014 for the non-transferable utility case, and Atakan 2006 for the transferable utility context).

the form of monetary transfers; in family economics, for instance, they typically affect the allocation of time between paid work, domestic work, and leisure; the choice between current and future consumption; or the structure of expenditures for private or public goods. Still, the possibility of utility transfers enables agents to negotiate, compromise, and ultimately exploit mutually beneficial solutions.

Not surprisingly, the existence of transfers between agents is one of the main factors structuring the matching literature. One polar framework (the so-called nontransferable utility—from now on NTU—model) assumes that transfers are not possible: there is simply no technology enabling agents to decrease their utility to the benefit of a potential partner. As a result, a match between two agents, say a and b, generates a gain for a and a gain for b, and these two gains are fixed; it is not possible for a to reduce her utility in order to increase b's.

In the alternative context, in which transfers are possible, a match creates a total surplus that has to be divided between agents; and an equilibrium must specify not only matching patterns—who is matched with whom—but also the supporting division of the surplus. In other words, the division, between partners, of the gain created by a match is *exogenous* in an NTU context (it is part of the definition of the matching problem), whereas it is *endogenous* when transfers are freely available.

When transfers are possible, a polar setting (based on transferable utility, from now on TU) postulates, moreover, that the transfer technology has a very striking property: it allows the transfer of utility between agents *at a constant "exchange rate,"* so that, for a well-chosen cardinalization of individual preferences, increasing my partner's utility by one "utile" (i.e., unit of utility) has a cost of exactly one utile for me, irrespective of the economic environment (prices, incomes, etc.). In that case, a given match generates some total gain, and moreover the division is such that individual utilities always *add up* to the total gain. Technically, the Pareto frontier, which represents the set of utility pairs that are just feasible given resource constraints, is a straight line with slope -1 irrespective of the constraints; the latter can only affect its intercept. Alternatively, a more general version (often called ITU for imperfectly transferable utility) allows for transfers, but recognizes that the exchange rate between individual utilities is not constant and is typically endogenous to the economic environment.

In practice, the choice of a particular model has to be based on its relevance for the context under consideration. In a pioneering study, Roth (1984) has studied the allocation of residents to hospitals; consecutive work has led to major improvements in the National Resident Matching Program (Roth and Peranson 1999). Similar advances have resulted in improved mechanisms for kidney exchange (Roth, Sönmez, and Ünver 2004) or the allocation of students to public schools (Abdulkadiroglu and Sönmez 2003). In these examples (and several others), transfers are excluded, either by regulation or because of the prevailing institutional constraints. In other cases, although transfers do take place, they are not endogenous to the matching mechanism, but determined outside of it; one may think, for instance, of the marriage market in society ruled by very rigid social norms. In all such cases, the relevant model belongs to the NTU family.[6] The large literature on the NTU matching game has been abundantly described, particularly in Roth and Sotomayor's excellent monograph (1990) and in the more recent survey of Hatfield and Kominers (2011).

The present book concentrates on the opposite situation, in which transfers are possible and endogenously determined (or at least constrained) by the equilibrium conditions. Many (if not most) real-life examples of matching involve transfers. Employers and workers match, and the resulting agreements involve wages, which are transfers from employers to employees; agents meet on markets to trade commodities and services, and these exchanges are based on prices and payments; individuals and households enter in risk-sharing agreements, which support transfers from the lucky to the unlucky; and so on. A prominent example is the marriage market, which will be abundantly discussed in this book. One can certainly write models in which intrahousehold transfers are not feasible; but these require very strong (and actually grossly counterfactual) assumptions. For instance, they must rule out the presence of private consumptions; otherwise, changes in the allocation of private consumptions between spouses de facto result in utility transfers. Even if all commodities are public—in the sense that they are always simultaneously consumed by both spouses—one must assume either that there exists a unique consumption good, or that household members have exactly the same preferences regarding their consumptions; if not, then again changing the consumption bundle varies the distribution of utility between spouses, which is formally equivalent to a transfer.

The main differences between NTU and TU (or ITU) frameworks will be discussed in detail in the next chapter. Let me illustrate it on a simple example. The introductory puzzles discussed above stress the importance of assortative matching—that is, whether people tend to marry their own kind, and why. Consider the simple case in which individuals differ by a single trait—say, physical

---

[6] Lastly, endogenous transfers typically require some minimum level of commitment between agents. A somewhat extreme example is provided by Lundberg and Pollak (2003), who consider a BIM (bargaining in marriage) framework in which no commitment is possible at all. In a BIM world, any promise I may make before marriage can (and therefore will) be reneged upon minutes after the ceremony; there is just no way spouses can commit beforehand on their future behavior. Moreover, "upfront" payments, whereby an individual transfers some money, commodities, or property rights to the potential spouse conditional on marriage, are also excluded. Then the intrahousehold allocation of welfare will be decided after marriage, irrespective of the commitment made before. Marriage decision will therefore take the outcome of this yet-to-come bargaining process as given, and we are back to an NTU setting in which each partner's share of the surplus is fixed and cannot be altered by transfers decided ex ante. This point will be discussed later (see section 3.5).

attractiveness. When do we expect beautiful people to marry beautiful people? The answer depends on preferences (obviously), but also on the nature of the matching game. In an NTU framework, one just needs to assume monotonicity of preferences: if everyone prefers a better-looking spouse, then at the equilibrium (the exact nature of which will be discussed later on) matching will be perfectly assortative on physical attractiveness. The intuition is clear: the best-looking woman will select the best-looking man, and he will accept; then the second-most-beautiful woman will choose the second-most-beautiful man (who happens to be the best looking among available men), who will accept, and so on.

When transfers are possible, however, the logic is quite different; in fact, it is easy to construct examples in which, although preferences are still monotonic (so that everyone still prefers a better looking spouse), the equilibrium matching is actually *negative assortative* (so that the worst-looking man marries the best-looking woman, and so on). To see why, note that the driving force behind realized matching is the bidding game implicit in the process. The "winner" is not the most attractive person, but simply the highest bidder; attractiveness does play a role, but can always be compensated for by a large enough transfer. The question that determines assortative matching, therefore, is the following: even though everyone would prefer a better-looking spouse (therefore willing to bid more to attract one), *how does this additional willingness to bid vary with the person's own attractiveness*? If less-good-looking men are more eager to match with a beautiful spouse (say, because they are richer, therefore more able to bid high), the outcome could be negative assortative.

Of course, whether beautiful people marry their own (or not) is not of much interest, at least for an economist. In the present book, I shall mostly concentrate on more explicitly economic traits, such as education, income, and human capital. But the basic logic will be similar: the drivers of assortative matching are fundamentally different in games with and without transfers. If anything, the mechanisms involved will be more complex in the presence of transfers, as we shall see in some details.

From an applied perspective, thus, the main distinction is between models that involve transfers and models that do not. Theorists, however, sometimes refer to a different classification. Namely, they distinguish TU models from all other models, thus gathering NTU and ITU frameworks within the same category—which, in a somewhat confusing way, is often referred to as NTU models. This distinction is justified by the very specific, theoretical status of the TU model; as we shall see below, it is the only one for which stability—the main equilibrium concept—is equivalent to surplus maximization. It follows that the conceptual tools used to analyze TU models are often very specific; for instance, there exists a close relationship between matching models and a class of mathematical problems known as *optimal transportation* (or *Monge-Kantorovitch*) problems, and this relationship is unique to the TU framework. On the contrary, both NTU and

ITU models often refer to a similar methodology, particularly to variants and extensions of the basic Gale-Shapley algorithm.[7]

Throughout this book, and for the sake of clarity, I shall nevertheless stick to the tripartite distinction NTU-TU-ITU. In particular, the NTU label will be exclusively applied to situations in which utility *cannot* be transferred between agents.

1.3 Matching and the Household

1.3.1 Household Behavior: Existing Models

As I said earlier, an important motivation for the study of matching models is their applications to family economics. It may therefore be useful to briefly summarize the state of the art in that field.

The Unitary Model

Historically, the most commonly used model of household behavior has been the static unitary model. The main assumption implicit in this approach is that households behave as single decision makers, independently of the number of household members. As a consequence, the unitary approach characterizes the decisions of married couples about consumption, labor supply, and household production in the same way it characterizes the decisions of people living on their own. This assumption is equivalent to postulating that the household's preferences can be represented using a unique utility function that does not depend on prices, incomes, or any exogenous factor, independently of the number of household members.

The unitary model is a natural starting point for modeling household behavior, since it makes the model tractable, simple to test, and easy to estimate. Whether it is a good description of household behavior is, however, a different question altogether. A strength of the unitary model is that it generates testable implications that can be used to answer that question. Thus, a well-known implication of the model is that the demand functions it generates have specific properties. One is that the corresponding Slutsky matrix (i.e., the matrix of *compensated* price effects on demand) should be symmetric and negative semidefinite. These properties have been tested and generally rejected.[8] A second testable implication is *income pooling*. In the unitary model, individual nonlabor incomes y^1 and y^2 impact household behavior only through the budget constraint and only as the

[7] See, for instance, subsection 3.1.9 below.

[8] See, for instance, Lewbel (1995), Browning and Chiappori (1998), Dauphin and Fortin (2001), Haag, Hoderlein, and Pendakur (2009), Dauphin and Fortin (2001), and Kapan (2010).

sum $y = y^1 + y^2$. As a consequence, after controlling for total nonlabor income y, individual nonlabor incomes y^1 and y^2 should not affect household decisions: what matters is total income, not the source of its different components. The income-pooling property has been thoroughly tested and generally rejected, since individual nonlabor income affects household behavior in ways that go beyond the effect of total income on budget constraint.

A possible reason for the rejection of income pooling is that the unitary model aggregates individual preferences in a way that is not consistent with the data. It is plausible that households make actual decisions by assigning higher weight to the preferences of members that are perceived to be more important or, equivalently, have more power within the household. The power of a person in a group is generally influenced by her or his outside options, which in turn depend on a collection of variables, such as individual income, wealth, wages, and human capital. If this is the case, households aggregate preferences in a way that depends on all those variables. In the unitary model, this possibility is ruled out since individual preferences can only be aggregated by using some fixed household index, which must be independent of any additional variable.

The Collective Model

The limitations of the unitary representation have become increasingly clear;[9] an alternative representation has emerged, usually referred to as the "collective model" (Chiappori 1988, 1992). Collective models of the household explicitly recognize that households generally consist of several individuals who may have distinct utilities. This recognition implies that the intrahousesold decision process plays a key role in determining behavior. Collective models assume that household decisions are efficient in the sense that they are always on the Pareto frontier. Remember that the *Pareto set* is defined as the set of utility pairs that are reachable within the household (through adequate transfers and coordination of decisions). The Pareto frontier is the frontier of that set; that is, if an outcome is located on the Pareto frontier, then it is not possible to further increase an agent's well-being without reducing that of someone else. Said otherwise, the efficiency assumption requires that, whatever decision is made, no alternative decision would have been preferred by *all* agents.

This axiomatic nature is a distinctive feature of collective models. They do not rely on specific assumptions on the way household members achieve an efficient outcome, such as Nash bargaining (which I will discuss later on). They simply assume Pareto efficiency, which is satisfied if for any decision the household makes, there is no alternative choice that would have been preferred by all household

[9] See, for instance, Alderman et al. (1995).

members. While the assumption of Pareto efficiency is undoubtedly restrictive, collective models are sufficiently general to include as special cases most of the static models used to study household behavior. A nonexhaustive list includes the unitary model, but also models based on cooperative game theory (for instance, Nash-bargaining models of household behavior, pioneered by Manser and Brown 1980 and McElroy and Horney 1981), including more recent formulations (for instance, the "separate sphere" model of Lundberg and Pollak 1993) or models based on a market equilibrium, as proposed by Grossbard-Shechtman (1993), Gersbach and Haller (1999), and Edlund and Korn (2002).

In practice, efficiency has a simple, technical translation: a decision is efficient if and only if it maximizes a weighted sum of members' utilities. The corresponding weights are called the Pareto weights; they may, in principle, depend on prices, individual income, and any variable that may affect individuals' respective bargaining positions. In particular, Pareto weights have a natural interpretation in terms of respective decision *powers*. The notion of power in households may be difficult to define formally. Still, it seems natural to expect that when two people bargain, a person's gain increases with the person's power. This somewhat hazy notion is captured very effectively by the Pareto weights: for any *given* cardinalization of utilities, the larger a person's Pareto weight, the larger that person's utility at the end of the decision process.

The collective theory of household behavior is by now well understood. In particular, we now have a set of necessary and sufficient conditions that fully characterize demand functions stemming from a collective framework (Chiappori and Ekeland 2006). Moreover, conditions have been derived under which individual preferences and the decision process can be recovered from the observation of household behavior (Chiappori and Ekeland 2009a,b); these conditions take the form of a "exclusion restrictions," since they require that, for each agent, there exists at least one commodity this agent does *not* consume (while other agents do).

A crucial property of the collective approach, however, is that it takes the household as given. In a typical, collective approach, both household composition and intrahousehold allocation of power are exogenously given; what the collective theory does is (i) derive necessary and sufficient conditions characterizing the demand functions generated by such a framework, and (ii) provide conditions under which the sole observation of household behavior is sufficient to *identify* the model—that is, to uniquely recover individual preferences and the decision process (as summarized by the Pareto weights associated with some specific cardinalizations of individual utilities). The next step, obviously, would be an "upstream" theory that would *endogenize* both household composition ("who marries whom?") and the resulting intrahousehold allocation of power. To reach that goal, two main paths have been followed by the literature: bargaining models and matching models. I will briefly discuss the first option; the remainder of this book will be devoted to the second.

Noncooperative Models

Before that, let me briefly mention an alternative family of models based on the notion that individuals do not always agree on the various decisions a household has to make, and that the resulting decision process is systematically conflictive, in the (strong) sense that spouses fail to cooperate even when cooperation would be beneficial for all. Specifically, while recognizing that the existence of public consumption within the household is a strong motivation for household formation, noncooperative models assume that spouses, when deciding how much they would be willing to spend on public expenditures, always disregard the benefits their partner would derive from these. Technically, individuals thus play a noncooperative game of private provision of the public goods. A precise description of these models is beyond the scope of this book; the interested reader may, for instance, refer to Browning, Chiappori, and Weiss (2014). Let me just mention two points. First, a recent literature has shown some surprising (and largely counterfactual) consequences of noncooperative models. For instance, whatever the number of public goods consumed by a two-person, noncooperative household, there can be at most one goods to which both spouses contribute; all other public goods are exclusively funded by one of the spouses. Moreover, whenever such a jointly contributed good exists, then household demand exhibits a strong version of the income-pooling pattern discussed above: a (small) redistribution of resources across spouses has no impact whatsoever on either public or even individual consumptions.

Second, very few papers in the matching literature actually use noncooperative models; besides their counterfactual predictions, such models do not seem fully consistent with the efficiency properties implied by TU matching. However, an interesting exception is provided by a recent article by Del Boca and Flinn (2014), who use marital sorting patterns as "out-of-sample" information to assess whether household behavior is efficient or not. Their approach develops a likelihood-based metric to compare marriage market fits under the two alternative behavioral assumptions of efficiency and inefficiency. Empirical estimation on a sample of households drawn from the Panel Study of Income Dynamics finds strong evidence supporting the view that household behavior is (constrained) efficient—a further validation of the collective approach.

1.3.2 Bargaining Models of the Household

A natural tool to formalize the endogenous genesis of individual powers within the household is cooperative game theory.[10] The approach, here, is axiomatic: given

[10] For a more detailed presentation, see Browning, Chiappori, and Weiss (2014).

the context of the game, theory provides a way to determine the outcome based on some properties that the solution concept must satisfy. The "context" is defined by two components: individual preferences on the one hand and *outside options* (or *threat points*) on the other hand—the latter representing the utility an individual would receive should no agreement be reached.

Bargaining models assume that the outcome of the decision process is Pareto-efficient. Clearly, if no point within the Pareto set can give the agents at least their threat points, then no agreement can be reached, since at least one member would lose by agreeing. In the opposite case, both agents can gain from the relationship; then an agreement will be reached, and the goal of bargaining models is to analyze how threat points influence the location of the chosen point on the Pareto frontier.

While bargaining models seem to provide a natural solution to our problem—how should we endogenize Pareto weights?—they still require two choices to be made: selecting a bargaining solution concept and defining the threat points.

Choosing the Bargaining Solution

I will start with the solution concepts. The most commonly used bargaining solution was proposed by John Nash in the early 1950s. Nash derived this solution as the unique outcome of a set of axioms that any "reasonable" solution must satisfy. Some of the axioms are uncontroversial (and, as a matter of fact, are shared by all commonly used bargaining solutions). One is individual rationality: an agent will never accept an agreement that is less favorable than his or her threat point. Another is Pareto efficiency, as discussed above. A third mild requirement is invariance with respect to affine transformations:[11] if both the utility and the threat point of an agent are transformed by the same increasing, affine mapping, the prediction about the equilibrium outcome of cooperation does not change. Note, however, that a nonlinear transform will change the outcome; that is, Nash bargaining requires a *cardinal* representation of preferences.

The last two axioms are more specific. One is symmetry; it states that if utilities and threat points are permuted between members, then the outcomes are simply switched. Natural as it may sound, this assumption may still sometimes be too strong. In many socioeconomic contexts, for instance, male and female roles are by no means symmetric. Fortunately, Nash bargaining can easily be extended to avoid the symmetry assumption.

The last and crucial axiom is independence. It can be stated as follows. Assume that the set of available opportunities (the Pareto set) shrinks, so that the new Pareto set is within the old one, but the initial equilibrium outcome is still feasible;

---

[11] An affine mapping is of the form $f(x) = ax + b$.

then the new equilibrium outcome will be the same as before. In other words, the fact that one member misses some opportunities that she had before does not affect her bargaining position with respect to the other member. This requirement alone implies that the Nash solution maximizes some function of the utilities of the two partners. Actually, Nash shows that the only outcome compatible with these axioms maximizes the product of individual "gains," the latter being defined as the difference between a person's utility when reaching an agreement and the person's threat point.

While Nash's bargaining concept is often used, it is by no means the only one. An alternative concept, proposed by Kalai and Smorodinsky (1975), replaces independence with the following *monotonicity* property. Consider two bargaining problems such that (i) the range of individually rational payoffs that player a can get is the same in the two problems, and (ii) for any given, individually rational utility level for player a, the maximum utility that player b can achieve (given the Pareto frontier) is never smaller in the second problem than in the first. Then player b should do at least as well in the second problem as in the first. In other words, if one enlarges the Pareto set by inflating b's opportunities while keeping a's constant, this change cannot harm b. Kalai and Smorodinsky prove that there exists a unique bargaining solution that satisfies all the previous axioms except for independence, which is replaced with monotonicity, and formally characterizes the corresponding solution.

Lastly, Nash himself suggested that one should provide noncooperative foundations to the bargaining solutions derived from axioms. The most influential framework is the model of Rubinstein (1982), in which players make alternating offers until one is accepted. When time matters through a constant discount factor, there exists a unique, subgame perfect equilibrium of this noncooperative game, which is characterized by the requirement that each player should be indifferent with regard to accepting the current offer or waiting for an additional round and making an offer that the opponent would accept. Binmore, Rubinstein, and Wolinsky (1986) have analyzed the link between these noncooperative formulations and the axiomatic approaches. Specifically, they study a model in which the bargaining process may, with some probability, be *exogenously* interrupted at each period. This model has a unique, subgame perfect equilibrium; moreover, if one allows the time interval between successive offers in both models to decrease to zero, then the equilibrium converges to the Nash bargaining solution.[12]

The main message of this brief review is that the choice of a bargaining solution concept is neither innocuous nor obvious; for instance, whether, when describing household decision processes, the independence axiom sounds more "realistic" than the monotonicity one is probably open to discussion.

[12] For a more complete discussion of two-person bargaining, see Myerson (1991, ch. 8).

Choosing the Threat Points

Next, how should one translate the abstract notion of "threat point" in the particular case we are considering—i.e., household bargaining? A first point is that the corresponding choice is crucial. Indeed, a result due to Chiappori, Donni, and Komunjer (2012) states that *any* Pareto-efficient allocation can be derived as the Nash bargaining solution for an ad hoc definition of the threat points. This implies that any additional information provided by the reference to bargaining theory must come from specific hypotheses on the threat points—that is, on what is meant by the sentence: "no agreement is reached."

Several ideas have been used in the literature. One is to refer to divorce as the "no agreement" situation. Then the threat point is defined as the maximum utility a person could reach after divorce. Such an idea seems well adapted when one is interested, say, in the effects of laws governing divorce on intrahousehold allocation. Another interesting illustration would be public policies such as single-parent benefits or the guaranteed employment programs that exist in some Indian states; Haddad and Kanbur (1992) convincingly argue that the main impact of the program was to change the opportunities available to the wife outside marriage, with potential consequences on intrahousehold inequality, even (and perhaps mostly) for couples who eventually did not divorce. However, choosing divorce as the threat point is probably less natural when minor decisions are at stake: deciding who will walk the dog is unlikely to involve threats of divorce.

A second idea relies on the presence of public goods, and the fact that noncooperative behavior typically leads to inefficient outcomes. The idea, then, is to take the noncooperative outcome as the threat point: in the absence of an agreement, both members provide the public good(s) egotistically, not taking into account the impact of their decision on the other member's welfare. This version captures the idea that the person who would suffer more from this lack of cooperation (the person who has the higher valuation for the public good) is likely to be more willing to compromise in order to reach an agreement. A variant was proposed by Lundberg and Pollak (1993), who postulate that each public good belongs to the "separate sphere" of one of the spouses, who, in the absence of an agreement, becomes the unique decision maker for the corresponding expenditures.

There is, however, something deeply unsatisfactory with any of the previous choices. Start with the model based on private provision of public goods (or its separate spheres variant): it only considers outside options *within* marriage. This is acceptable for minor issues, much less so when considering consequential choices. Divorce is a crucial aspect of family economics, if only because nearly half of all marriages end in divorce. It seems unwise to disregard it: among the "outside options" a spouse is likely to consider in case of a serious disagreement on a very important issue, switching to a new partner can hardly be dismissed prima facie. Granted, divorce threats are unlikely to determine who will walk the dog. Then again, who will walk the dog is not the most fascinating problem economists would want to understand. And when it comes to really substantial decisions—fertility, choice of a job, choice of a location, etc—divorce issues just cannot be easily omitted.

What about, then, models adopting divorce as a threat point? The problem here is that these models typically adopt an exclusively local approach, when a global perspective would be needed. In the end, what matters is not the possibility of divorce by itself, but the utility each spouse can expect to have after divorce (this is the technical definition of a threat point in this context). But the latter, in turn, depends on many aspects: the probability of remarriage, the likely characteristics of the new spouse, and also the (expected) distribution of powers within the new couple. The last point is obviously problematic: it implies that in order to predict the distribution of powers within any given couple, one needs to know the distribution of powers in all other couples, including future and potential ones.

For an economist, a situation of this type is by no means unfamiliar. After all, we cannot predict the wage structure that will prevail in a given firm without knowing the wage distribution in the rest of the economy, precisely because this distribution impacts the outside options of both parties. In other words, these are market-wide phenomena, and they can only be addressed at the market level—the “global” perspective that I was alluding to.

The problem is especially serious when we want to address issues that, by nature, are global. Assume, for instance, that we want to model the impact on intrahousehold decision processes of a reform of the legislation governing divorce. For any given couple, the reform will modify the spouses’ outside options—i.e., utilities in case of divorce—in a number of ways. For instance, it will typically affect actual divorce decisions; but this impacts what could be called the “market for remarriage” by changing the number and profiles of divorcees looking for a new partner. In turn, such changes alter all aspects of the costs and benefits of divorce—the remarriage probability, the set of potential new spouses, and ultimately the distribution of powers. In this context, an analysis of the consequences of the reform can only be performed at a global level: what is needed is a characterization of the new equilibrium that may emerge after the reform.

To summarize: if one is to take seriously the idea that divorce should be considered as a reference threat point for modeling bargaining within marriage—as I think one should—then only a “general equilibrium” model of matching on the marriage market can be expected to generate the desired outcome. This intuition is by no means new. The abstract of the seminal, 1973 paper by Gary Becker, published in the *Journal of Political Economy*, starts with the following two sentences:

> I present in this paper the skeleton of a theory of marriage. The two basic assumptions are that each person tries to do as well as possible and that the "marriage market" is in equilibrium.

Becker's work was obviously imperfect (although it amounted to much more than a "skeleton"). The model he used was far from general; some of the simplifying assumptions he made would be frowned upon today. Matching theory has made considerable progress since then, both on theoretical and empirical fronts; it is only fair to say that the technicalities of modern approaches exceed (and sometimes contradict) those of Becker's initial contribution. But for all these advances, the main conclusions have essentially confirmed Becker's vision, and particularly the central intuition of his 1973 paper—namely, that the division of surplus (or of power) between spouses could be derived from the nature of the marriage market equilibrium. The remainder of this book will illustrate how such a derivation can work.

1.4 Content

Large numbers of contributions have been devoted to matching theory in recent decades. Matching has become a huge field, and I will certainly not try to cover all of it in this short book. In particular, and as the title of the book indicates, I will exclusively consider matching models in which transfers between agents are possible. Therefore, I will not cover the literature on matching with NTU. Several reasons motivate this choice. One is that the two families of models—with and without transfers—are, in many respects, quite different; many of the intuitions that emerge from an NTU framework would not be robust to the introduction of transfers. Another reason is that my primary interest is family economics, a context in which transfers are usually crucial; analyzing household decisions under the maintained assumption that there is no way a husband and a wife can transfer utility between them, although technically possible, does not sound too promising. Last but not least, the literature on NTU matching has already been covered, notably by Roth and Sotomayor's (1990) excellent monograph, as well as in several subsequent surveys (for instance, Hatfield and Kominers 2011).

A second limitation is that I will exclusively analyze bipartite (or one-to-one) matching models, i.e., models in which agents from two distinct populations (say, men and women) match by pair. In particular, I will not consider frameworks involving either many-to-one or many-to-many matching. Again, these are less relevant for the type of applications I will consider, including the marriage market. Also, these models raise specific problems (for instance, a stable matching may not always exist) that would be outside the scope of this short essay. The cost of this choice is that many interesting questions will be largely left aside. For instance, I will not spend much time discussing the links between matching and auction or general equilibrium theory; even in the field of family economics, I will not consider issues such as polygamy, although some of the results could be extended to this question.

Throughout the book, I will mostly adopt what could be called an "applied theory" perspective. That is, although I will present and discuss the main theoretical results underlying the approach, I will mostly consider how these abstract concepts can be used to represent and think about real-life issues; the reader interested in more technical aspects can refer to the recent monograph by Galichon (2015). Similarly, I will not discuss in depth the technical issues related to the econometric implementation of matching models; again, the interested reader can refer to existing surveys such as Graham (2011) and Chiappori and Salanié (2016). Moreover, matching models with transfers can involve either finite or continuous distributions of agents. The main results apply to both settings. However, I will put a particular emphasis on continuous models, in which agents' characteristics are drawn from an absolutely continuous and atomless distribution; for many applications, this context is both richer and more tractable.

Lastly, I will mention a few applications of matching models to such real-life issues as the empowerment effect of abortion legalization or the surge in women's demand for higher education. The choice is purely idiosyncratic, and reflects both my own interests regarding economic issues and my unavoidably incomplete knowledge of the large number of models that have been developed over the last two decades. As I said earlier, matching with transfers is a booming field, which I could not dream of covering in this short essay. If, however, some readers are convinced of both the importance of the issues at stake and the practical usefulness of matching models to address them, then this book will have achieved its goal.

Index