NBER WORKING PAPER SERIES

STRUCTURAL MODELS:
INCEPTION AND FRONTIER

Sebastian Galiani
Juan Pantano

Working Paper 28698
http://www.nber.org/papers/w28698

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2021, Revised January 2022

This manuscript is being prepared as a chapter in the methods section of the forthcoming Handbook of Labor, Human Resources and Population Economics. We thank Jerome Adda, Philipp Eisenhauer, John Ham, Alvin Murphy, David Neumark, Ken Wolpin and Chenyu Yang for their comments on an earlier version of this manuscript. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Structural Models: Inception and Frontier
Sebastian Galiani and Juan Pantano
NBER Working Paper No. 28698
April 2021, Revised January 2022
JEL No. C01,C52

ABSTRACT

We discuss the past, present and future of the structural approach in empirical microeconomics, starting with its inception in the 1970s and 1980s. Our focus is on the use of the structural approach in labor economics, broadly defined to include population economics, human capital and related fields. In the hopes to reach a wider audience that might not be as familiar with the pillars of the structural approach, we first provide an overview of well-known features, setting the stage for a more up-to-date discussion of current developments. We discuss how to identify the need for a structural model, and key steps involved in how to formulate one. We also discuss issues of identification and estimation and highlight advantages and disadvantages of this approach, including the controversial issue of external validity. We then describe the current frontier of this approach, which increasingly reflects integration efforts with “design-based” strategies. This integration provides opportunities to both, validate structural models and enhance the credibility of their identification. We highlight why, whenever possible, it is best to pursue both of these goals, reserving some of the credible exogenous variation for identification and some for validation. While quasi-experimental variation can be useful in pursuit of both of these goals, we discuss why RCTs provide a first best opportunity in terms of out-of-sample validation. We conclude with thoughts about the future of the structural approach.

Sebastian Galiani
Department of Economics
University of Maryland
3105 Tydings Hall
College Park, MD 20742
and NBER
sgaliani@umd.edu

Juan Pantano
Eller College of Management
University of Arizona
1130 East Helen Street
Tucscon, AZ 85721
jpanta.uchicago@gmail.com

1 Introduction and Epistemological Background

The structural econometric approach has a long tradition in applied microeconomics. In this approach, the underlying economic model is made explicit and the econometric analysis is inextricably intertwined with this model. While there are various definitions of "structure", all of which have their merits, in this chapter we take "structure" to mean, in the spirit of Marschak (1953) and Hurwicz (1962), invariance to environmental changes. That is, in the structural approach, the goal of estimation is to recover model features that are structural in the sense of stability under environmental manipulation. The value of the approach is predicated on the idea that once these policy invariant or "deep" features that govern decision-making under alternative environments are uncovered, one can more confidently use the structural model to evaluate how the decision-makers' behaviors change when facing alternative environments, including new ones not encompassed by historical variation.

Despite its undeniable promise, the structural approach remains relatively underused in various fields in applied microeconomics due to both, its complexity and its apparent reliance on a large number of assumptions relative to other approaches that impose less structure on the inference problem. As argued by Angrist and Pischke (2010), in recent years a credibility revolution has had a large influence in the practice of empirical microeconomics, with an emphasis on transparent sources of identification and less reliance on fully specified economic models. Indeed, the last four decades witnessed two diverging approaches to empirical work in the applied microeconomics fields: the so-called "structural" and "reduced-form" approaches. These classifications are not perfect and there have been attempts to relabel these approaches as "model-based" and "design-based", a taxonomy that is more suitable. For simplicity however, we continue the custom of referring to "structural" models as those in which the object of estimation are the primitive parameters of a fully specified model of behavior.

This chapter focuses on the use of structural models in empirical microeconomics, with particular emphasis in labor economics, including models of human capital accumulation, and related fields such as population and family economics. The chapter complements the views of Wolpin (1996), Meghir (2006), Keane (2010a), Keane (2010b), Rust (2010), Wolpin (2013), Low and Meghir (2017) and Blundell (2017) on the general value and usefulness of the structural approach in empirical microeconomics, while also being candid about some of its limitations, as emphasized in Angrist and Pischke (2010) and Imbens (2010). For the most part we focus on the "traditional" structural approach, where the optimizing unit, be it an individual, a couple or a firm, is assumed to make decisions as if they had successfully solved the optimization problem that the econometrician assumes them to solve. The two key assumptions are then a) that the goal that the decision-making unit is trying to reach is correctly assumed and b) that, whatever that goal is, the unit is able to successfully come up with the best choice in pursuit of that goal. It is fair to say that these assumptions have come under increased scrutiny, even by early pioneers of the structural approach.[1] Our review is centered on the traditional paradigm because we feel it is still a very valuable approach in many contexts and that its credibility can be improved. Further, we feel it is important to understand how to overcome the limitations the structural approach may have even when those powerful identifying assumptions being maintained are true, before venturing into even less restricted models of human behavior. Despite this traditional focus, we briefly touch on and provide pointers to more recent work that calls for a relaxation of these assumptions and proposals for more integration with insights from behavioral economics into the models estimated using the structural approach.

Our review hopes to provide a less technical and perhaps more balanced introduction to the structural approach for the novice practitioner that wants to become familiar with the approach. While there are several opinion statements on its value, we are not aware of comprehensive textbooks or reference work that covers all the technical material that is relevant for a thorough understanding of the structural approach.[2] Rather, the relevant body of knowledge seems to be scattered through a dizzying array of specialized surveys. We hope to fill this gap by providing as broad an overview as possible and a methodological road-map for this approach. We also provide key references to structural work specifically applied to the themes of this Handbook.

The rest of the chapter is organized as follows: in Section 2 we provide key references to the literature that uses estimable structural models in labor economics and closely related areas such as human capital and population economics. Section 3 describes the structural approach in detail, discussing formulation, identification, estimation and validation of structural models. Next, Section 4 highlights advantages and disadvantages of structural models. In Section 5 we layout how, in our view, the credibility of the structural approach can be further enhanced. In this section we highlight a recent trend towards a systematic integration with the type of clear sources of exogenous variation that are the hallmark of modern non-structural approaches. We go on to argue that this integration should be pursued for both identification and validation of these models, if one is to maximize the credibility of the

[1] Contrast for example, the increasingly nuanced tone of Rust (2014, 2019) with his more enthusiastic stance, just a few years earlier, on the value of the traditional structural approach in his Rust (2010) review of Keane (2010b).

[2] Some of the material in the textbooks by Adda and Cooper (2003) and Christensen and Kiefer (2009) do cover some of the fundamentals concepts relevant for dynamic structural estimation as applied to the themes of this chapter, but their scope is somewhat different as they also devote much of their content to macroeconomics and finance perspectives

structural approach. We provide some concluding remarks in Section 6 with some thoughts about the future of structural models.

2 Structural Models in Labor, Human Resources and Population Economics

This chapter is primarily concerned with a methodological description of the structural approach as applied in the fields which are the substantive focus of this Handbook. Due to space limitations, it is not meant to provide a comprehensive, exhaustive catalog of the literature that has used structural methods in labor, human resources and population economics. In this section and throughout the chapter we provide references to seminal contributions and/or to work representing the current frontier in these fields. Whenever possible we also provide pointers to specialized surveys that help in tracking the evolution of the structural approach in these areas. While we limit ourselves to structural work most relevant to the themes in this Handbook, it is important to recognize that some important methodological insights, relevant for the structural approach that we do not cover here, have been developed in other fields.[3]

The origins of the structural approach in labor economics can be traced back to Heckman (1974) who modeled female labor supply at the intensive and extensive margins as a result of an explicit, internally consistent utility maximization problem.[4] In the last 45 years this literature has developed in multiple dimensions, integrating savings, uncertainty, on-the-job training, occupational choice, job search, matching, turnover and joint spousal decisions about labor supply. Several specialized surveys exists that describe these developments (e.g.

[3]In particular, we do not focus on Industrial Organization (IO), a field where the structural approach has been particularly fruitful and we refer the reader to Reiss and Wolak (2007) and Ackerberg et al. (2007) for general surveys of the structural approach in that field and Athey and Haile (2007) and Gentry et al. (2018) who review the structural estimation of auction models, an area whose insights are being exported into more general models of asymmetric information. Einav and Levin (2010) and Nevo and Whinston (2010) discuss the success of the structural approach in the field of IO, and attempt explanations as to why it gained more widespread adoption in IO relative to the fields we focus on. The economics of education and health economics have also begun to rely increasingly on structural models. For example Einav and Finkelstein (2018) discuss how a structural model allows to go beyond the exogenous experimental or quasi-experimental variation in the data when analyzing ex-post moral hazard in the demand for health care. However, they caution that different structural models might be consistent with this variation. This is related to the question of structural model identification, which we discuss below. Similarly, Ferreyra (2007) shows the value of the structural approach to analyze the effect of voucher policies, a classic question in the economics of education.

[4]Some argue that Heckman (1974) is not strictly a fully structural approach, and classify it instead as semi-structural because he doesn't work with the marginal rate of substitution (MRS) between consumption and labor as derived from a specific utility function, but rather takes a linear approximation to log(MRS).

Eckstein and Wolpin (1990), Wolpin (1995), Blundell and Macurdy (1999), Blundell et al. (2007), and Keane et al. (2011) and provide detailed summaries of these richer models of labor supply and a history of the evolution of the structural approach in labor economics. Structural models of labor supply have also been extensively used to understand the consequences of different tax (see e.g. Keane (2011, 2015)) welfare (see e.g. Chan and Moffitt (2018)) and retirement (see e.g. French and Jones (2017)) policies.

The structural approach is also used in population economics and in models of human capital formation. Wolpin (2003), Heckman et al. (2006) and Belzil (2007) provide surveys of the structural literature that aims to understand schooling decisions and the returns to schooling. Building on the early work of Willis and Rosen (1979), Keane and Wolpin (1997) provided an important extension where schooling decisions are the result of a dynamic, sequential process by forward-looking individuals that is integrated with labor supply and occupational choices.[5] In the same vein, Eckstein and Wolpin (1999) provide a structural analysis of high school dropout decisions. In addition to the methodological contributions, a key substantive message of this literature was that there is a large degree of heterogeneity in skills accumulated by the the mid- to late teenage years when this important, life defining schooling decisions (e.g. high school dropout or college attendance) must be made. This has spurred interest in estimating the structural parameters of the technology of skill formation for cognitive and non-cognitive skills (Cunha and Heckman (2008), Cunha et al. (2010), Agostinelli and Wiswall (2020)), particularly for developmental stages earlier in life, ranging from birth until the teenage years. A more recent literature integrates these technologies into fully structural models of household decision-making which explicitly consider parental preferences and derive the household choice of optimal inputs allocated for children's skill formation, as in the work of Del Boca et al. (2014).

Turning to structural work in population economics, beginning with Wolpin (1984), a growing literature uses estimable dynamic structural models to investigate fertility decisions. Wolpin (1997) and Hotz et al. (1997) provide early surveys of this literature and Adda et al. (2017) provides one of the most recent applications, integrating fertility with occupational choice to investigate the career cost of children. Following Choo and Siow (2006), increasingly rich models of the marriage market have been estimated to understand patterns of marriage formation in a friction-less framework. Building on the collective framework of Chiappori (1988) a growing literature estimates structural models of intra-household alloca-

[5]The methods developed by Keane and Wolpin (1994, 1997) for estimating more realistic dynamic structural models with larger state spaces and choice sets led to a surge of applications. Keane and Wolpin (2009), Todd and Wolpin (2010) and Keane et al. (2011) provide surveys of applications that use dynamic structural models (particularly, dynamic programming models of discrete choice) in labor economics and other applied microeconomic fields.

tion under the assumption that allocations among household members are Pareto efficient. An alternative approach (e.g. Del Boca and Flinn (2012)) allows for the possibility of inefficient allocations in a non-cooperative framework. Chiappori and Mazzoco (2017) provide a comprehensive summary of this line of research. A synthesis of the structural literature on marriage market equilibrium and intrahousehold allocation is emerging (see Chiappori et al. (2019), Gousse et al. (2019) and Shephard (2019)).

Finally, another literature uses estimable structural models to understand human resources questions, such as optimal compensation of workers and managers within firms. When effort is unobserved and cannot easily be inferred from output, a principal-agent problem arises. A large body of work in personnel economics summarized in Prendergast (1999) provides key theoretical insights into this problem and in particular, the form of the optimal compensation contracts in those settings. Examples of structural approaches that build on these models include Ferrall and Shearer (1999) and Paarsch and Shearer (2000) for worker compensation and Margiotta and Miller (2000) and Gayle et al. (2015) for managerial compensation. In Section 5 we discuss recent work by Misra and Nair (2011) and d'Haultfoeuille and Février (2020) where exogenous variation in compensation contracts is used to either validate or estimate the structural model.

3 The Structural Approach

There is no general recipe on how these models can be formulated and estimated. We describe some generalities but the model and the estimation strategy is something that is chosen on a case-by-case basis depending on the question of interest and the available data. In this section we review what is a structural model, discuss how to assess the need for one, and review how such a model is formulated, identified, estimated, validated.

3.1 Structural Models Defined

What is a structural model? At a very abstract level we follow Matzkin (2007) and define a a structural model as given by

$$\mathcal{M}(Y, X, \varepsilon; F, G) = 0 \tag{1}$$

Where Y denotes an observed vector of endogenous variables (e.g. choices, outcomes) and (X, ε) are vectors of observable and unobservable variables that are neither choices nor outcomes, but allow for heterogeneity that is relevant for decision-making. F denotes a vector of (possibly non-parametric) functions and G denotes the vector of (possibly non-parametric) model distributions for random variables. $\mathcal{M}() = 0$ denotes a vector of structural relations between the primitive objects in the model.[6]

While the above definition is quite general it might be too abstract to keep the discussion concise, so we follow Wolpin (2013) in using a simple model of labor supply to focus on the essential issues that generally come up in the structural approach and to illustrate the basic ideas throughout the chapter. The model is extremely simple and in no way intends to be portrayed as frontier structural work in labor supply. We again point to the surveys of the literature referenced in Section 2 for those interested in the richer models. The basic labor supply model is an ideal pillar to describe the structural approach as some other topics like, fertility, human capital accumulation, investment in children or choice of compensation with unobserved work effort all can be seen as extensions of this basic model.

Consider the simplest, prototypical labor supply model. Suppose individual i has standard preferences for consumption c_i and leisure l_i. She has some non-labor income I. She can use this income to purchase a consumption good. She can also choose to supplement this income by working a number of hours h_i, and thus sacrificing some leisure to enhance her consumption opportunities. To decide how much to work she solves:

$$\begin{aligned} &\max_{\{c_i,l_i,h_i\}} U(c_i, l_i, X_i, \varepsilon_i^u) \\ &\quad \text{s.t. } \; c_i = w_i h_i + I_i \\ &\qquad\quad T_i = h_i + l_i \\ &\qquad\quad w_i = \exp(\alpha_X^w X_i + \varepsilon_i^w) \end{aligned} \tag{2}$$

where c_i denotes consumption, l_i denotes leisure, T_i is the total time available to individual i (e.g. 24 hours per day, 365 days per year), w_i represent an hourly wage, X_i denote a vector of exogenous individual characteristics (e.g. education etc) that in general enters both, the utility function and the wage equation. $\{\varepsilon_i^U\}$ and $\{\varepsilon_i^W\}$ represent unobserved (to the econometrician) distaste for work and productivity that affects wages, respectively. In general, $\{\varepsilon_i^U\}$ and $\{\varepsilon_i^W\}$ might be correlated in the population under study. Their joint distribution is given by $g(\varepsilon_i^U, \varepsilon_i^W)$

This model can be rewritten in terms of the optimal choice of h only. Once an optimal choice of h is found, consumption and leisure can be directly obtained from the budget

[6] An even more general formulation would include Υ, the unobserved vector of endogenous choices. These unobserved choices are often ignored in applied structural work relying on separability assumptions about how these choices enter the structural relations in the model.

constraint:

$$\begin{aligned} \max_{\{h_i\}} & U(w_i h_i + I_i, T_i - h_i, X_i, \varepsilon_i^u) \\ \text{s.t. } & w_i = \exp(\alpha_X^w X_i + \varepsilon_i^w) \end{aligned} \tag{3}$$

We denote $h^* = h(w, I, X, \varepsilon^u)$ the labor supply function that is the solution to this problem and assume for the moment that everyone chooses to work some amount, perhaps because their non-labor incomes is too low to support their desired consumption. If wages and non-labor income are exogenous, the function $h(w, I, X, \varepsilon^u)$ is called a reduced form because it describes an endogenous variable as a function of the exogenous variables in the model. Note that this simple structural model of labor supply fits the general framework described in (1). Table 1 provides a crosswalk between the general formulation and our simple labor supply example.

Table 1: The Structure of a Simple Labor Supply Model

General Structure	Labor Supply Model
X:	$\{X, I\}$
ε:	$\{\varepsilon_i^u, \varepsilon_i^w\}$
Y:	choices:$\{h, l, c\}$, outcomes:$\{w\}$
F:	U(, , ,)
G:	$g(\varepsilon_i^u, \varepsilon_i^w, X, I)$
$\mathcal{M}() = 0$:	$h_i - \arg\max_h \{U(w_i h + I_i, T_i - h, X_i, \varepsilon_i^u)\} = 0$
	$w_i - \exp(\alpha_x^w X_i + \varepsilon_i^w) = 0$
	$T_i - [l_i + h_i] = 0$
	$w_i h_i + I_i - c_i = 0$

It is important in any particular application to understand how the specific model at hand fits this general structure.

3.2 Assessing the Need for a Fully Structural Approach

Our focus in this review is on the full specification of a structural model, on the premise that identification of all its primitive features is necessary to answer the research question of interest. In many cases, however, estimation of a fully specified structural model is far from necessary to answer the research question of interest.

So it is important then to first ask ourselves, whether we need to recover the full structural model and identify early on when we can do just fine with approaches that rely on less structure. As emphasized by Matzkin (2007), if the question of interest resides in the labor supply function one does not need to uncover the full structure. For example, if we are interested in the effect of wages on hours of work we can rely on observed exogenous variation in wages to identify their effects on hours of work, without need of recovering first primitives like $U(,,,)$. Estimating $\frac{\partial h()}{\partial w}$ will suffice in that case. However, as noted early on by Burtless and Hausman (1978), if we would like to understand how labor supply would react to the introduction of a labor income tax $\tau(w, H)$ with complex brackets, marginal rates and notches it might be necessary to work with a fully structural model. This is particularly true in the hypothetical case that the tax is to be introduced for the first time.[7].

We endorse the idea that simpler approaches that directly rely on experimental or quasi-experimental variation and do not impose the additional structure might be preferable in those cases. There is also a middle ground in which only some structural features or combinations of them are sufficient to answer the research question. In that case, only the structure that is actually necessary to identify those features should be maintained and imposed on the data. Similarly, Heckman (2010) focuses on the value of maintaining only the assumptions that are necessary to answer a class of research questions, where this partial knowledge of the structure is all that is needed. He uses a generalized Roy model as an example. He notes that under this minimalist approach one may not need to recover all of the deep structural parameters of the Roy model. Rather, combinations of them such as the marginal treatment effect (MTE) profile may suffice. The MTE is the average treatment effect for individuals at a given quantile of the distribution of unobserved resistance to treatment. The MTE profile, often reported graphically, traces out this effect across the full range of of unobserved resistance to treatment. Policy-relevant treatment effects (PRTEs) can then be computed by integrating the MTE profile with appropriate MTE weights that the individuals induced to take treatment, not by the available instrument but by the policy of interest.[8] This is an example of the so-called Marschak's Maxim after Marschak (1953). The summary statistic approach of Chetty (2009) builds on similar ideas, focusing on the minimal structure needed. It should be noted, though, that this so-called semi-structural approach can only be used to analyze the effects of a class of policies as long as the variation induced by these policies is encompassed by the existing variation used to estimate the model. While many policies fit this class, not all of them do.

Another type of "middle ground" approach sometimes referred to as "semi-structural". It involves fully specifying the dynamic economic model as in the fully structural approach.

[7]Wolpin (2013) discusses how one might do without a fully structural approach if there is such historical variation in taxation that can be leveraged. See also Ichimura and Taber (2002) for related ideas.

[8]For a less technical introduction to Roy models and the MTE concept see Cornelissen et al. (2016) and the chapter by Navarro (2022) in this Handbook.

However, when it comes to estimation in this "semi-structural" approach one can, in a discrete choice model, replace the value functions associated with a given choice with a linear or polynomial function of the relevant state variables. This avoids the need to computationally solve the dynamic problem to obtain the exact form of that continuation value. While this is not a fully structural approach in the sense that one can no longer separate current utility from the expected future value, some research questions can still be answered in this "semi-structural" way, as shown in, for example, in Blau and Gilleskie (2001).

Our focus here is then on recovering the full set of structural parameters, with the understanding that interest resides in questions which are only possible to answer with such a deeper knowledge. As emphasized by Wolpin (1996) the need for a fully structural approach becomes apparent when the question of interest, represented in the structural model as change in the exogenous variables or a known modification of the structure, is not encompassed by existing variation in the data. In those cases, a model is needed to go beyond the variation in the data and analyze, ex-ante, the causal effects of new policies that have never been experienced. One view is that in principle then, structural inference would seem to escape the dictum of "no causation without manipulation" associated with the Rubin causal model as defined in Holland (1986). The structural approach uses auxiliary manipulations to learn the deep invariant principles that govern behavior. It then leverages those to conduct causal inference about hypothetical manipulations that were never experienced. We do not weigh in on these semantic debates but, more constructively, focus on providing a review of the structural approach for those who are willing to pay the price of maintaining structural assumptions in exchange for additional knowledge. A more novel focus of our review is to emphasize ways in which the credibility of this additional knowledge can be improved.

3.3 Formulating Structural Models

While the above representation attempts to be general enough to encompass every type of structural model, we feel it might be worth working through our specific labor supply example to discuss the idiosyncrasies of the different types of structural models. In other words, how is a structural model formulated? What decisions must be taken? In this subsection we provide a step-by-step guide through many of the decisions that must be taken when fully specifying a structural model.

1. **Agents.** A first decision is whether to have a single-agent or a multiple-agent model. We focus on single-agent models where it suffices to characterize the behavior of an isolated individual. However, there are settings in which the decisions of more than one individual are linked and each individual has its own utility function. For example,

in the context of the labor supply model above one may want to consider the case in which two spouses must decide how much each one works. In that setting, it might be necessary to consider the distinct utilities functions of each of the spouses, the distinct wages they face and whether the goods they consume and provide utility are private (each spouse consumes its own) or public (there is joint consumption). In addition, one must make an assumption about how spouses interact to reach a decision. One possibility, as discussed below, is to impose the assumption of Pareto efficiency and work out the restrictions that it implies for household members' consumption and leisure allocations. In special cases, efficiency may be the solution to some bargaining game between the spouses, but this is not necessary. Alternatively one can rely on non-cooperative game-theoretic models which may or may not deliver Pareto efficient outcomes. The structure that needs to be identified now include the two distinct utility functions of the household members and the sharing rule used to decentralize a planner's solution to the household problem. Some restrictions on how the utility functions of the spouses depend on the other spouse's allocations are generally necessary to identify the model.

2. **Equilibrium.** Another important modeling choice is regarding equilibrium considerations. Is the setting best described by a model where a single side (demand or supply) of a single market is modeled? Or is it more appropriate to use an equilibrium model (where demand and supply in a single market are modeled) or even a general equilibrium model where demand and supply in all inter-related markets are modeled. Most structural work in empirical microeconomics abstracts away from equilibrium considerations and focuses instead on more detailed microeconomic modeling that captures institutional detail and heterogeneity with more granularity. Most work in structural microeconometrics features the added value of zooming in and better capture first order effects that come from a more realistic, less stylized modeling of decision-making, at the price of ignoring aggregate consistency, especially when general equilibrium effects are thought to be of second order in magnitude. However, beginning with Heckman et al. (1998) a handful of papers have developed strategies to estimate empirically grounded, micro-econometric structural models of labor supply that account for equilibrium considerations.[9] The job search literature, described below has also embraced the equilibrium approach. The marriage literature referenced in Section 2 also estimates equilibrium models of the marriage market. In the context of the simple labor supply model above one would need to specify the distribution of agents in the economy

[9]See, among others, Lee (2005), Lee and Wolpin (2006), Meghir (2006), Lee and Wolpin (2009), Johnson and Keane (2013) and Lise et al. (2015).

and construct an aggregate labor supply function by summing all the individual labor supplies $H(w) = \sum_i h_i(w_i)$ and an aggregate labor demand $D(w)$ function that specifies how much labor the production side of the economy demands at various wages. One can then note that observed wages must be those that equate aggregate demand and supply.

$$H(w) = D(w)$$

The advantage of an equilibrium framework is most apparent when large environmental changes are analyzed. For example, if a significant tax or transfer reduces incentives to work, and aggregate labor supply no longer meets demand at original wages, then the equilibrium restriction will ensure that wages increase, and the partial equilibrium reduction in hours of work will be mitigated to some extent. The equilibrium approach does not necessarily require estimating additional structural parameters other than the technology parameters that underlie the aggregate labor demand function.

3. **Planning Horizon and Forward Looking Behavior.** Are agents myopic? or are they forward-looking? Do agents take into account the future consequences of current actions? Is the time horizon that agents entertain in the model finite or infinite? Once the model is forward looking a new structural parameter, the discount factor becomes part of the structure. This parameter controls how the individual values future utility in terms of current utility. Once the agents are forward looking and take into account future consequences of current actions many extensions to the basic model are possible. As discussed in Rust (1994), the solution methods for forward-looking models are quite different depending on whether the horizon is finite or infinite. In the finite case, models are solved through backwards recursion. Infinite horizon models are typically solved using fixed point methods that recover the value functions. We discuss some extensions to the basic model in (2) that open up once individuals are forward-looking.

   (a) Savings: once the model has more than one period and the individual is forward-looking, he might be interested in saving or borrowing as in MaCurdy (1981) to transfer income from a period in which the wage is high to a period in which it is low. We consider a simple 2-period extension of the model in (2) where now the individual must choose labor supply in both periods and how much to save or borrow, s_i, in the first period. The same ideas apply in a multi-period model with 45 or 50 periods where a period is a year. For simplicity we assume he no longer has exogenous non-labor income $I_{i,t}$ in any period.

$$
\begin{aligned}
\max_{\{h_{i,1},h_{i,2},s_i\}} \; & U(c_{i,1}, l_{i,1}, X_i, \varepsilon_i^u) + \delta U(c_{i,2}, l_{i,2}, X_i, \varepsilon_{i,2}^u) \\
\text{s.t. } & c_{i,1} = w_{i,1} h_{i,1} + s_i \\
& c_{i,2} = w_{i,2} h_{i,2} + s_i(1+r) \\
& T_{i,t} = h_{i,t} + l_{i,t} \text{ for } t = 1,2 \\
& w_{i,t} = \exp(\alpha_x^w X_i + \varepsilon_{i,t}^w) \text{ for } t = 1,2
\end{aligned} \tag{4}
$$

Notice that if the utility function in both periods is the same (the function itself is not indexed by period, only its input arguments are) no additional structure is added. The relevant interest rate, r, must be added to the set of exogenous observable variables X_i, but a common value is typically assumed. Then the only new structural parameter that is added to the structure is δ, the intertemporal discount factor. One also needs to decide whether the distribution of exogenous unobservable variables $(\varepsilon_{i,2}^u, \varepsilon_{i,2}^w)$ is the same in both periods. In general, one could let this distribution and/or the utility function for the second period to be different, for example, if the anticipated shocks to tastes for leisure or the wage shocks are expected to have different variance in the population or the person anticipates to enjoy each hour of leisure more in the second period than in the first.

(b) Uncertainty: Once dynamics are allowed, the individual might be uncertain about the value that some of the variables may adopt in the future. A crucial distinction opens up in the dynamic model on whether the heterogeneity that realizes in period 2 $(\varepsilon_{i,2}^u, \varepsilon_{i,2}^w)$ reflects something that was known and anticipated by the individuals in period 1 or, instead, reflects unanticipated shocks. In the latter case, for example, one could extend the model in (4) by letting the individual be uncertain about the values of $(\varepsilon_{i,2}^u, \varepsilon_{i,2}^w)$. In that case, the second term in the objective function would be

$$\delta E\left[U(c_{i,2}, l_{i,2}, X_i, \varepsilon_{i,2}^u)\right] \tag{5}$$

Where $E[\cdot]$ is the mathematical expectation operator and the expectation is taken with respect to the joint distribution of $(\varepsilon_{i,2}^u, \varepsilon_{i,2}^w)$. In models with uncertainty then a new feature added to the structure is the belief that individuals have about the joint density $(\varepsilon^u_{i,2}, \varepsilon^w_{i,2})$. A common approach in the traditional structural approach is assume that these beliefs coincide with the actual joint density that is already part of the stochastic structure G in the 2-period model. More recent work uses data on beliefs, relaxing the assumption that they coincide with realized distributions. Finally if the asset in which the person saves is risky, an additional source of uncertainty would be with respect to the asset returns r. Again, one would need to add to the structure some beliefs about the distribution of this risky asset and provide assumptions and/or data to identify them.

(c) Learning by Doing: individuals may recognize that working today not only provides labor income today but increases the wage in the future as they become more productive at the job the more they work, as in Imai and Keane (2004). In that case we only need to specify how hours of work in period 1, $h_{i,1}$ increase the wage in period 2, $w_{i,2}$. For example, we may assume that each additional hour of work in period 1 increases the wage by α^w_h percent in period 2.

$$w_{i,2} = \exp(\alpha^w_X X_i + \alpha^w_h h_{i,1} + \varepsilon^w_{i,2}) \tag{6}$$

A new structural parameter α^w, capturing returns to experience or learning by doing is added to the model structure that needs to be identified. These models are particularly useful to understand wage growth and the costs of career interruption. As argued by Keane (2011), they are also important to recover the right labor supply elasticities. In models where different types of jobs are available, it is also important to understand mobility across jobs, depending on whether the experience accumulated in one job can be transferred to another occupation in period 2 or if it is instead specific to the occupation chosen in period 1.

(d) Learning about oneself: Even if there is no learning by doing, the individual may not know how productive in the job she is and this might be important if her wage depends on her productivity. She learns about this by observing signals and using these signals for updating her priors. A key for learning is that there is no deterministic relationship between her effort and her output, otherwise she would learn immediately. This may give her incentives to experiment so as to accelerate learning about parameters she is uncertain about. Miller (1984) pioneered the structural estimation of this type of learning models in the context of occupational choice.[10] To see the essential point in the context of the simple dynamic model in (5), let us assume first that X is just a constant and the worker is paid according to her productivity. The individual is uncertain about the value of α^w and whenever she works one hour her output is variable for random reasons η^w that she cannot track down. For each hour h she works during the first period she is paid $w_{i,1,h} = \alpha^w + \eta_{i,1,h} + \varepsilon^w_{i,t}$ and there are no savings. So we must now include an expectation operator in the first period, not because she does not know $(\varepsilon^u_{i,1}, \varepsilon^w_{i,1})$, but because she is uncertain about α^w. Since she is differentially productive each hour, she is paid a different hourly wage each hour, $w_{i,1,h}$. We characterize her beliefs about α^w with a density $g_1(\alpha^w)$ and note that the corresponding density for the updated belief in period 2, $g_2(g_1, h_{i,1}, \{w_{i,1,h}\}_{h=1}^{h=h_{i,1}}, g_\eta(\eta))$, depends on the initial belief g_1, the number of hours of work $h_{i,1}$, the value of productivity signals she obtained $\{w_{i,1,h}\}_{h=1}^{h=h-i,1}$ and the distribution of the noise g_η. These signals are used to update her initial beliefs. Therefore, while the constraints are the same, the objective function becomes:

$$
\begin{aligned}
\max_{\{h_{i,1}, h_{i,2}\}} \; & E_{\alpha^w}\left[U(c_{i,1}, l_{i,1}, X_i, \varepsilon^u_i)\right] + \delta E_{\varepsilon,\alpha^w}\left[U(c_{i,2}, l_{i,2}, X_i, \varepsilon^u_{i,2})\right] \\
\text{s.t. } & c_{i,t} = w_{i,t} h_{i,t} \text{ for } t = 1, 2 \\
& T_{i,t} = h_{i,t} + l_{i,t} \text{ for } t = 1, 2 \\
& w_{i,t} = \exp\left(\sum_h \left[\alpha^w + \eta_{i,t,h}\right] + \varepsilon^w_{i,t}\right) \text{ for } t = 1, 2
\end{aligned}
\tag{7}
$$

where the expectation in the first period is only over α^w and the expectation in the second period includes, in addition, the expectation over $(\varepsilon^u_{i,2}, \varepsilon^w_{i,2})$. Note that in terms of added structure, we must now consider the following features: a) the initial period's belief g_1, b) the distribution of the unobservable productivity shocks g_η and c) the updating procedure that she uses to update her initial beliefs upon observation of signals from her first period. It is common to impose strong distributional assumptions for a) and b) and directly postulate a particular belief updating procedure (e.g. Bayesian). These models are particularly useful to understand firm tenure as it is often the case that employers and employees are uncertain about how productive the employee will be in the new job.

[10] Structural estimation of learning models has also been used in population economics and models of human capital accumulation. For example, Brien et al. (2006) model learning about match quality among cohabiting partners and spouses whereas Mira (2007) models learning about child mortality risk in a dynamic model of fertility decisions. Arcidiacono (2004) models student learning about their own ability to succeed in college in a model of college decision-making.

(e) Job Search: consider a multi-period ($T > 2$) extension of the model in (4). Suppose there are different opportunities to work at different wages, but the individual may not receive a job opportunity in every period. In some periods then, she cannot work and needs to rely on some exogenous non-labor income (or on her savings). Furthermore, once she receives an opportunity, if she accepts it, she might not be able to search for alternative, better opportunities. Therefore, as long as there are better potential opportunities, she may not accept offers when they are received and instead wait for a better one. The model is now then given by:

$$\begin{aligned}
&\max_{\{h_{i,t}\}} E\left[\sum_{t=1}^{T} \delta^t U(c_{i,t}, l_{i,t}, X_i, \varepsilon^u_{i,t})\right] \\
&\quad \text{s.t. } c_{i,t} = w_{i,t}h_{i,t} + I_{i,t} \text{ for } t = 1...T \leq \infty \\
&\quad\quad T_{i,t} = h_{i,t} + l_{i,t} \text{ for } t = 1...T \leq \infty \\
&\quad\quad w_{i,t} = \exp(\alpha^w_x X_i + \varepsilon^w_{i,t}) \text{ for } t = 1...T \leq \infty \\
&\quad\quad \pi_t = \Pr(\text{offer at t}) \text{ where } 0 < \pi_t < 1
\end{aligned} \tag{8}$$

where the expectation is now not only over the $(\varepsilon^u_{i,t}, \varepsilon^w_{i,t})$ that may realize in any period, but also over the probability that a job offer is actually available in each period, π_t. Note that π_t could in general be a function of features of the individual's history up to time t. Note that the job offer probabilities π_t are the new structural feature that must be considered.

Structural models of job search are particularly useful to analyze alternative unemployment benefit policies by letting the non-labor income $I_{i,t}$ to be augmented by unemployment benefits in periods when the individual does not receive a job offer or chooses not to accept one.

Job search models have been extended in many dimensions and have been used to address many other questions beyond the issue of unemployment benefits. They can be broadly categorized by whether they are set up in discrete or continuous time and by whether the environment is stationary (infinite horizon and all relevant parameters of the job search environment do not vary over time or non-stationary (where there is a finite horizon and/or at least one of the parameters varies over time). Flinn and Heckman (1982) were the first to structurally estimate the infinite horizon stationary job search model in continuous time. Wolpin (1987) was the first to estimate a finite horizon model in discrete time where the non-stationarity arises from the finite horizon and by allowing of duration dependence in job offers. van den Berg (1990) first estimated the general non-stationary job search model in continuous time. The job search literature has been perhaps one of the most successful at implementing the structural approach in labor economics. Several specialized surveys by Eckstein and Wolpin (1990), Devine and Kiefer (1991), Wolpin (1995), Canals and Stern (2002), Postel Vinay and Robin (2006), Eckstein and van den Berg (2007), Christensen and Kiefer (2009) and Flinn (2010) provide guidance to this voluminous literature. Recent work by Taber and Vejlin (2020) integrates the structural job search and self-selection literatures.

As it may have become apparent, in any dynamic model, regardless of its details, we will have to consider the discount factor δ. The identification of the discount factor δ is often challenging so it is common in applications to fix it to a conventional value depending on the time unit of the model. In some cases, however, this parameter is estimated, and even allowed to be heterogeneous in the population (e.g. Sauer (2004), Arcidiacono et al. (2007), French and Jones (2011)). Frederick et al. (2002) provide an early summary of papers estimating discount factors. Recent work allows for time-inconsistency in dynamic optimization and goes beyond the time-consistent, exponential discounting paradigm (e.g. Fang and Silverman (2009), Fang and Wang (2015), Chan (2017), Mahajan et al. (2020))

4. **Time Unit.** When the model is dynamic, another key modeling choice is whether to set up the model in continuous time or in discrete time. In case the time is chosen to be discrete, the proper frequency in calendar time must be specified. Are periods equivalent to a year? a month? a week? Similarly, when setting up the model in continuous time one must specify if the choices can be taken at each instant in real time or, despite the continuous time formulation, opportunities to make choices arise only at certain times. With the exception of the job search literature described above, most structural work uses a discrete time framework. Arcidiacono et al. (2016) and Abbring (2012) discuss estimation of structural models in continuous time.

5. **Choice Set.** Another important decision is whether to model the choices that are available to the agents in the model as discrete or continuous. As pointed out by Miller (1999), there are tradeoffs involved in the modelling of continuous versus discrete choices. In some cases the choice variable is naturally discrete. When the variable is continuous, though, one must trade-off measurement error problems that are more

common among measures of a continuous choice and the more stringent assumptions often needed to identify discrete choice models. While continuous choices are more common in macroeconomics, dynamic structural work in empirical microeconomics often emphasize discrete choices, following the seminal contributions by Gotz and McCall (1984), Wolpin (1984), Miller (1984), Pakes (1986) and Rust (1987) and subsequent innovations proposed by Hotz and Miller (1993), Hotz et al. (1994), Keane and Wolpin (1994), Aguirregabiria and Mira (2002), Su and Judd (2012) and Arcidiacono and Miller (2011).[11] Recent methods have been adapted to account for a mixture of discrete and continuous choices. See Iskhakov et al. (2017) and Blundell et al. (2016).

To emphasize the key difference in a discrete choice model, we can return to the static model with continuous choice of work hours in (3) and restrict it to be just a model of binary choice, where the individual decides on a simple choice d regarding whether to work $d_i = 1$ (that is $h_i > 0$) or not work $d_i = 0$ (that is $h_i = 0$). That is:

$$\begin{gathered}\max_{d_i \in \{0,1\}} U(w_i d_i + I_i, d_i, X_i, \varepsilon_i^u) \\ \text{where } w_i = \exp(\alpha_X X_i + \varepsilon_i^w)\end{gathered} \tag{9}$$

The individual observes ε_i^w and therefore knows her hourly wage. She also knows her non-labor income I and observes her taste for leisure (ε_i^u). Given its discrete nature, the model is now characterized by the particular value ε_i^{u*} that solves:

$$U(w_i + I_i, 1, X_i, \varepsilon_i^u) = U(I_i, 0, X_i, \varepsilon_i^u) \tag{10}$$

There exists $\varepsilon_i^u = \varepsilon_i^{u*}$ such that the individual is indifferent between working and not working. Those with $\varepsilon_i^u > \varepsilon_i^{u*}$ choose not to work and those with $\varepsilon_i^u < \varepsilon_i^{u*}$ choose to work. This model can be easily extended to a dynamic setting where the individual makes discrete choices over 2 or more periods in the same way that we discussed above for the model with continuous choice of hours.

While we emphasize the opportunities to formulate structural models with either continuous and discrete choices, it is important to mention that the discrete choice approach is often preferred. Many choices that are the object of interest in empirical microeconomics are discrete and, for those that are inherently continuous, researchers

[11] We refer to the surveys of structural microeconometric methods for estimation of dynamic programming models of discrete choice provided in Eckstein and Wolpin (1989b), Rust (1994) Aguirregabiria and Mira (2010), Keane et al. (2011) and Arcidiacono and Ellickson (2011).

often choose to discretize them anyway. A reason for this is that continuous choices generally lead to continuous variables in the state space. This makes the model solution more costly in the context of dynamic models, as interpolations become necessary. Similarly, to derive policy functions in the context of continuous choice models with closed form solution it might be necessary to interpolate. Models where both choices and states are discrete are typically simpler to code and discrete choice models can also leverage a convenient distributional assumption on the structural error terms. As shown by Rust (1987), if the error terms associated with the discrete choices are distributed Type 1 Extreme Value, then the expected maximum of the alternative specific value functions has closed form and the choice probabilities have the familiar multinomial logit form.

6. **State Space.** In all models, whether static or dynamic, it is particularly important to be explicit about what is known to the decision maker at the time she makes choices because uncertainty about things that will realize after a choice needs to be taken into account. When the model is dynamic, the sequence representation of the problem is often reformulated using a dynamic programming representation, using a Bellman Equation. For example, the model in (8) can be re-written as follows:

$$\begin{aligned} V_t(\Omega_{i,t}) = \max_{h_{i,t}} E\left[U(c_{i,t}, l_{i,t}, X_i, \varepsilon^u_{i,t}) + \delta^t V_{t+1}(\Omega_{i,t+1})\right] \\ \text{s.t. } c_{i,t} = w_{i,t}h_{i,t} + I_{i,t} \\ T_{i,t} = h_{i,t} + l_{i,t} \\ w_{i,t} = \exp(\alpha^w_x X_i + \varepsilon^w_{i,t}) \\ \pi_t = \Pr(\text{offer at t}|\Omega_{i,t}) \text{ where } 0 < \pi_t < 1 \end{aligned} \tag{11}$$

where $\Omega_{i,t}$ denotes the state space, which includes everything that is known and relevant for the decision-maker to make her choices. The value function at time t is given by $V_t(\Omega_{i,t})$. It is a function of the state variables and captures the expected present discounted value of the remaining utility under the restriction that not only the current choice is optimal but also future behavior in periods other that t will also be selected optimally. It is very important to keep track of the state space. These are variables that are important for decision-making in the sense that either directly affect utility in the current period or they affect the distribution of the state variables in the next period. Two key distinctions on how the state variables evolve are : a) deterministic versus stochastic evolution and b) exogenous versus endogenous evolution. Some state variables like age evolve over time deterministically and exogenously. Previous work experience is a common state variable in dynamic models of labor supply, and also evolves deterministically, but in an endogenous way, as it reflects past work decisions. The number of children would be an endogenous and stochastic state variable in a dynamic model of labor supply and fertility choice that features imperfect fertility control. One could also have state variables like weather, that evolve exogenously and stochastically and may affect choices.

In dynamic models then, an additional feature of the structure are the laws of motion for the stochastic state variables:

$$\Pr(\Omega_{i,t+1}|\Omega_{i,t}, d_{i,t}) \tag{12}$$

7. **Objective Function.** What is the utility function that agents maximize? We have been using a standard expected utility formulation where the utility in the various periods is additively separable. This is certainly the most commonly adopted approach. However, one could in principle make other choices here, with regards to how individuals in the model deal with uncertainty by using non-expected utility theories. One could also argue that individuals do not aim to maximize the discounted sum of future utilities. Most of the structural work in empirical microeconomics is firmly grounded on use of expected utility and time-consistent dynamic optimization with exponential discounting. As we discuss below, however, new paradigms are emerging, drawing on insights from behavioral economics, promoting the use of non-expected utility theories to deal with uncertainty and time-inconsistent approaches to deal with forward-looking dynamics.

8. **Observability from Decision Maker's and Econometrician's perspectives.** What, choices, states and outcomes are observable to the agent in the model and what choices, states and outcomes are observable to the econometrician? In most cases, although not always, the agent in the model observes more than the econometrician. For example, the individual deciding whether to work or not in model (9) observes everything but the econometrician does not observe ε_i^u.[12]

9. **Observed Heterogeneity.** In principle any structural features of the model can be allowed to vary based on observable variables. Are the primitive structural objects in the model the same across decision-makers with different observable characteristics?

[12] Note that while the econometrician does not observe ε_i^w directly, he can back it out given the observation of wages, the parameters of the wage equation and the fact that $\varepsilon_i^w = w_i - \alpha_X^w X_i$.

The answer is most certainly no. In the models above we already let U and w vary with X. One could similarly let g_ε,δ,π and $\Pr(\Omega_{i,t+1}|\Omega_{i,t}, d_{i,t})$ to vary with X. In the case of dynamic models this could be particularly costly as different dynamic programming problem must be solved for individuals with different structure. In general one must be judicious in what type of observable variables to allow so as to keep the computational cost under control. However, particular care must be chosen when allowing some features of the model and not others to vary by X. Restricting some of the features to not vary by X often amounts to imposing exclusion restrictions that may turn out to be critical for identification, and so they need to be carefully justified.

10. **Unobserved Heterogeneity.** It is often the case that even after allowing for some observed heterogeneity one may suspect that there is still some residual heterogeneity in structural features that is unobserved by the econometrician. For example, in the context of labor supply model above one could argue that the utility function is different across people in the sense that different individuals are willing to trade-off consumption and leisure at different rates. By allowing for unobserved tastes for leisure (ε_i^u) the static model already allowed for this. However, in a dynamic context one may want to distinguish between a permanent component of (ε_i^u), call it μ^u that is fixed, time-invariant and known to the individual and a separate component ν_t that captures unanticipated shocks to the taste for leisure. That is, one may want to have $\varepsilon_{i,t}^u = \mu_i^u + \nu_{i,t}$. This separation between uncertainty and heterogeneity has important implications for estimation. One then needs to specify the stochastic structure for μ_i^u and $\nu_{i,t}$. While $\nu_{i,t}$ is often specified similarly to $\varepsilon_{i,t}^u$, using a continuous density $g(\nu_{i,t})$, μ_i^u is most often specified using a discrete distribution of so-called "types". Each individual has multinomial probability $\Pr(k)$ of belonging to each of these types $k = 1, ..., K$. This multinomial probability (the values that μ_k adopts in its domain and the probability mass at each of those values) then becomes part of the additional structure that must be identified and estimated in the fully structural approach.

11. **Functional Form.** While ideally one would proceed without imposing any functional form assumptions on the structural functions of the model (e.g. $U()$) this is almost never done. First, most often the available data is not sizable enough to allow for a fully non-parametric treatment of every feature. Moreover, even with infinite data, the structural model is often used to extrapolate outside the support of the data, requiring anyways functional form assumptions for such extrapolation. For example, in the labor supply model above one could, following Imai and Keane (2004), choose the following

functional form for the utility function

$$U(c_i, l_i, X_i, \varepsilon_i^u) = \frac{c_{i,t}^{1+\eta}}{1+\eta} - \exp(\psi X_i + \varepsilon_i^u)\frac{[T_{i,t} - l_{i,t}]^{1+\gamma}}{1+\gamma} \qquad (13)$$
$$\text{where } \eta < 0, \gamma > 0$$

where ε^u represent unobserved distaste for work.

12. **Distributional Assumptions.** Ideally one would not need to specify any assumption on the stochastic structure G, allowing it to be non-parametric. In practice, however, structural econometricians proceed by specifying distributional assumptions for the stochastic structure of the model. For example one could assume that $(\varepsilon^u, \varepsilon^w)$ is distributed bivariate normal

$$\begin{pmatrix} \varepsilon^u \\ \varepsilon^w \end{pmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_u^2 & \rho\sigma_u\sigma_w \\ \rho\sigma_u\sigma_w & \sigma_w^2 \end{pmatrix}\right]$$

Note that, importantly, when $\rho \neq 0$, wages will be endogenous in the labor supply function in the sense that $E[\varepsilon^u | w_i] \neq 0$

3.4 Investigating the Identification of Structural Models

The question of model identification is whether, in the context of model $\mathcal{M}(\cdot) = 0$, there is a unique structure (F, G) that generates the joint distribution of observable variables $\Pr(Y, X)$. A model is not identified if one can find two structures (F', G') and (F'', G'') that generate the same $\Pr(Y, X)$. Ideally, researchers should provide a formal proof of identification in the sense of Matzkin (2007, 2013) by mathematically proving that the structural parameters or features of interest (G, F) are uniquely recovered from the joint distribution of available data $\Pr(Y, X)$, the structure of the model and any auxiliary identifying assumptions. The question of identification asks whether there is a unique structure such that starting from the distribution of exogenous variables (X, ε) one can recover the observed distribution of endogenous variables given the exogenous variables $\Pr(Y|X)$. It is beyond the scope of this review to provide a detailed discussion of identification for all types of structural models. We direct the reader to French and Taber (2011) who provide identification results using Roy models of self-selection as an organizing principle and to Chiappori and Mazzoco (2017) who discuss identification of models of intra-household allocation. Early work by Keane (1992) notes the fragility of identification in simple static multinomial probit models, an important result for structural models with multiple discrete choices. Much progress has been made in the last 25 years on the identification of dynamic structural models. Important results were established in Rust (1994), Taber (2000), Magnac and Thesmar (2002), Kasahara and Shimotsu (2009) and Hu and Shum (2012) for dynamic discrete choice models. Abbring (2010) provides a survey of these early results but this remains an active area of research.[13] A recent literature (see, for example, Norets and Tang (2014)) has pointed out that not all counterfactuals of interest might be identified under some of the normalizations typically adopted for identification in these models. On the other hand, as argued by Kalouptsidi et al. (2021), even if a model is not non-parametrically identified, the counterfactual of interest might be.

While Matzkin (2007) provides an elegant formalization of the identification problem, in practice it is often difficult to prove identification formally given the complex, highly nonlinear structure of these models. Therefore, heuristic/intuitive arguments are often provided. Eckstein et al. (2019) provide an example of such arguments.

Sometimes it is argued that small standard errors are indicative of local identification. Indeed, when estimation proceeds via maximum likelihood, standard errors are often computed using the inverse of the information matrix. If the standard errors are small, that means that the likelihood function declines quickly when moving away from its maximum. Therefore, there are in the vicinity of the maximum, no alternative sets of parameters that could deliver the same likelihood. Similarly, when used a moment-based strategy, local identification fails if the Jacobian of the moment vector is not invertible (See Adda and Dustmann (2021)). Another informal strategy to provide evidence of local identification is available for those who structure their estimation strategy using moments. In that case, one can show whether and how all the moments change in response to small changes (one-at-a-time) in the structural parameters. If a parameter change results in no changes whatsoever in any moment, then that parameter is not identified. See for example, Adda and Dustmann (2021). This is useful and provides some reassurance in many contexts, but in general all the moments jointly identify all the parameters so exploring one parameter at a time does not provide a proof of global identification. Related ideas on this informal "sensitivity analysis" approach to investigating the empirical identification of structural models are discussed in Andrews et al. (2017, 2020a,b), Honoré et al. (2020), Jørgensen (2021) An alternative approach to identification is to show, via Monte Carlo exercises designed to replicate the empirical setting, that the proposed estimation strategy recovers the true parameters, starting from very

[13]See Blevins (2014), Bajari et al. (2016), Arcidiacono and Miller (2020), Abbring and Daljord (2020)) and Levy and Schiraldi (2020) for more recent developments.

different initial guesses.

3.5 Estimating Structural Models

Structural models are estimated in many different ways depending on details of the model and the data at hand. Moreover, the same structural model can be estimated in different ways. A comprehensive treatment of estimation details is beyond the scope of this chapter. In this section we provide some general points about estimation that come up often when taking structural models to the data. We limit ourselves to parametric structural models where the structural features in F have been specified using parametric functional forms and particular distributional assumptions have been made about the stochastic structure in G. Estimation therefore is limited to recovering the values of a vector of parameters $\theta = \{\theta^F, \theta^G\}$. where θ^F denotes the vector of parameters in the structural functions of the model and θ^G refers to the parameters of the distributions of random variables in the model, including the assumed distribution of unobserved heterogeneity, if any.

Like other models, structural models of microeconomic behavior make predictions about what the value of the endogenous variables Y_i is expected to be given observed values of the exogenous variables X_i. There are, broadly speaking, two approaches a researcher can pursue when taking the model to the microdata for estimation. Both approaches attempt to rationalize why individuals with the same observed X_i end up being observed with different choices and outcomes Y_i. One option is to assume that the model is correct and the reason why individuals have different values of Y_i is because the data was measured incorrectly. This is the measurement error approach, which essentially assumes that the model is right and the data is wrong. Any discrepancies between model predictions and what is observed reflects measurement error. A second approach that tends to be more popular is to assume that there is structural unobserved heterogeneity in the model. For example, in the model above the term ε^u represent heterogeneity in distaste for work that the econometrician cannot observe, but the individual is fully aware of. Individuals with the same X_i make different choices Y_i because of the different ε^u they have. We focus our review on the unobserved heterogeneity approach to estimation of structural models but mention briefly the measurement error approach as well.

3.5.1 Classical Approaches to Estimation

These models are often estimated by Maximum Likelihood or Generalized Method of Moments (GMM).[14] Since these methods are thoroughly discussed in standard econometrics textbooks, we provide a minimal discussion here to highlight how structural models are embedded within these estimation methods. Given the highly nonlinear nature of structural models, these estimation methods rarely have a closed form for the vector of structural parameters and therefore the estimation procedure must iteratively search for these parameters until the estimation criteria is optimized.[15] Most structural models that allow for some form of forward-looking behavior require solving dynamic programming models. When game-theoretic models are involved, some form of equilibrium must be considered and solved. These optimizing and/or equilibrium solutions that must be found at each trial of the structural parameter vector often require heavy computations, slowing down the estimation routine.

Maximum Likelihood. Maximum Likelihood estimation proceeds by constructing the likelihood function and then maximizing it with respect to the structural parameters θ. The likelihood function is nothing more than the joint density (or probability, in the case of discrete variables) of the observed data, $\Pr(Y, X, \theta)$. The likelihood function takes the realized data as fixed, and only the parameters in the vector θ are the arguments over which it is maximized, so we write it $L(\theta; Y, X)$. The logarithm of the likelihood function (i.e. the log-likelihood) is often maximized, as it is numerically more convenient.

In most cases the cross-sectional data is independent across individuals so the likelihood function is just the product across observations of the individual likelihood contributions, $L_i(\theta; Y_i, X_i)$

$$L(\theta; Y, X) = \prod_{i=1}^{N} L_i(\theta; Y_i, X_i) \tag{14}$$

Note that, $f(y_i, x_i; \theta) = f(y_i|x_i; \theta)f(x_i)$ and the distribution of the observable exogenous variables $f(x_i)$ does not depend on θ because that is not a feature that is explained by the model. Therefore one can focus solely on the density of y given x.

In the context of the labor supply model in (2) we have $y_i = \{h_i, \log(w_i)\}$ and $X_i = \{I_i, edu_i\}$. To estimate the model using cross-sectional data $\{h_i, \log(w_i), I_i, edu_i\}_{i=1}^{N}$ one would proceed by constructing the likelihood contribution for each observation. This is really just the joint density $f(h_i, \log(w_i)|I_i, edu_i; \theta)$. Note also that this is a simple model

[14] We limit our review to frequentist approaches. For Bayesian approaches to estimation of dynamic structural models of discrete choice see Geweke and Keane (2001), Norets (2009) and Imai et al. (2009)

[15] In very special cases, some of the structural parameters might be estimated in closed form using simple IV or two-way fixed effects methods. See for example MaCurdy (1981) and Blundell et al. (1998)

where we assume everyone works a positive number of hours and therefore we observe the wage offers for everyone. The likelihood contribution for observation i is then given by

$$L_i(\theta; Y_i, X_i) = f(h_i, \log(w_i)|I_i, edu_i) \tag{15}$$

An important step in the construction of the likelihood contribution is then to connect the primitive stochastic structure G with the densities of the endogenous variables. In our labor supply example this means linking the assumed density for $(\varepsilon^w, \varepsilon^u)$ with the densities of $(h, \log(w))$. For log wages, this is $f(\log(w_i)|edu_i)$ and given $\log(w_i) = \alpha_x^w X_i + \varepsilon_{i,t}^w$, then one can simply use a change-of-variable technique to recast $f(\log(w_i)|edu_i)$ as the density of the implied wage heterogeneity term ε^w.[16]

The contribution from observed work hours is derived similarly. Taking first order condition with respect to h in model (2) and using a particular utility function would lead to an optimality condition for hours of work. One can then recast the density of the observed data of hours of work in terms of the density for the structural unobserved heterogeneity in unobserved distaste for work.[17]

Maximum likelihood estimation of the discrete choice model in (9) proceeds similarly. In the discrete case, when building a likelihood contribution one must integrate over the distribution of possible values of the unobserved heterogeneity and come up with model-based probability for d rather than a density for h.

Consider, for simplicity, the utility linear function

$$\begin{aligned} U(C_i, d_i, X_i, \varepsilon_i^u) &= C_i - \psi_i d_i - \psi_1 C_i d_i \\ \text{where } \psi_i &= \psi_0 + \psi_x X_i + \varepsilon_i^u \end{aligned} \tag{16}$$

Using the same budget constraint as in (9), then the utility when working ($d_i = 1$) is just $U(C_i, 1, X_i, \varepsilon_i^u) = (1 - \psi_1)(w_i + I_i) - \psi_0 - \psi_x X_i - \varepsilon_i^u$. When not working it is just $U(C_i, 0, X_i, \varepsilon_i^u) = I_i$. Armed with the utilities from the two choices, in structural models of discrete choice it is common to then search for the value of the unobservable ε_i^{u*} that renders the person indifferent between taking any of discrete choices. In this case $\varepsilon_i^{u*} = w_i - \psi_0 - \psi_1 w_i I_i - \psi_x X_i$. For values of $\varepsilon_i^u \leq \varepsilon_i^{u*}$, the unobserved distaste for work is low enough and the individual works, whereas for $\varepsilon_i^u > \varepsilon_i^{u*}$, she chooses not to work. The finding of ε_i^{u*} and considering the inequalities around it are the analogous step to taking

[16] Note that the Jacobian for this simple change of variables is just $|1|$ because $\log(w_i)$ is additively separable in ε^w.

[17] It is often necessary to use change of variable techniques with complex Jacobians to accomplish this when the hours of work is not additively separable in the unobserved distaste for work.

first order conditions in a model with continuous choice. Note that for estimation purposes and to focus on the essence, we continue to assume that we observe w for all individuals, including those who choose not to work.[18] Further, assume that not only non-labor income I_i, but also wages, w_i are exogenous. Then, the likelihood contribution for observation i is given by

$$\begin{aligned} L_i(\theta; Y_i, X_i) &= \Pr(d_i | w_i, I_i, edu_i) \\ &= \Pr(\varepsilon_i^u \leq \varepsilon_i^{u*} | w_i, I_i, edu_i)^{d_i} \Pr(\varepsilon_i^u > \varepsilon_i^{u*} | w_i, I_i, edu_i)^{1-d_i} \\ &= \Phi\left(\frac{\varepsilon_i^{u*}}{\sigma_u}\right)^{d_i} \left[1 - \Phi\left(\frac{\varepsilon_i^{u*}}{\sigma_u}\right)\right]^{1-d_i} \end{aligned} \tag{17}$$

where $\Phi(\cdot)$ is the CDF of the standard normal distribution. The above likelihood contribution can be easily generalized to more realistic cases in which wages are endogenous (e.g. there is correlation between ε_u and ε_w) and accepted wage offers are only observed for those working. In that case one would again need to add the density of wages for those who work and consider how the conditioning on a given value of ε_w affects the density of ε_u.[19]

In dynamic models, ignoring permanent unobserved heterogeneity and assuming that the errors are independent over time, the likelihood contribution would simply involve the product over time of period-specific choice probabilities like the one in Equation 17. In a dynamic model one would need to find the cutoff values ε_{it}^{u*} by solving out the choice-specific value functions and considering their difference. In a finite horizon model one obtains the value functions by backwards recursion, starting from the last period and recursively solving period-specific values until the first period of the model. For infinite horizon models one typically finds the value function as the solution to the fixed point implicit in a Bellman equation like the one in Equation 11. For more details on the mechanical implementation of these general solution algorithms for dynamic programming problems in both, finite and infinite horizon models, see Rust (1996). Keane et al. (2011) provide details more focused on the backwards recursion procedure for finite horizon models common in Labor Economics.

Moment-Based Estimation An alternative to Maximum Likelihood estimation are

[18] Most work in labor economics, relaxes this assumption and takes into account the econometric problem of only observing wages for those who work.

[19] We refer the reader to the more specialized discussion of Maximum Likelihood available in several econometrics textbooks for details on the mechanics of how to maximize the function and how to compute standard errors. Note however that while computation for standard errors for structural parameters is standard, the computation of standard errors for counterfactual objects might not be. In many cases the delta method is not appropriate. Ham and Woutersen (2011) and Eisenhauer et al. (2021) investigate this in detail and propose solutions.

moment-based estimation strategies such as the Method of Moments (MOM) or the Generalized Method of Moments (GMM). This approach tends to use less information than Maximum Likelihood, but it is often computationally simpler and, in some models may offer the opportunity to avoid the need for making certain distributional assumptions. The moment approach considers the moment condition

$$E[m(h, w, I, X; \theta^*)] = 0 \tag{18}$$

where θ^* is the true vector of structural parameters. If only one moment condition is available from the model and there is more than one parameter to estimate, it is possible to expand the set of moment conditions by using a vector of exogenous instruments Z such that $E[m(\cdot)|Z] = 0$ and $E[Zm(h, w, I, X; \theta^*)] = 0$ by the law of iterated expectations. For example if in the model above (w, X, I) are assumed to be exogenous then $Z = [w, X, I]'$.[20]

If $\text{Dim}(Zm(\cdot)) = \text{Dim}(\theta)$ the model is just identified and estimation proceeds by methods of moments, numerically finding the vector θ that satisfies

$$\frac{1}{N}\sum_{i=1}^{N} m(h, w, I, X; \theta) = 0 \tag{19}$$

If $\text{Dim}(Zm(\cdot)) > \text{Dim}(\theta)$ one can use GMM, where

$$\hat{\theta} = \underset{\{\theta\}}{\arg\max} \left(\left[\frac{1}{N}\sum_{i=1}^{N} m(h, w, I, X; \theta) \right] W \left[\frac{1}{N}\sum_{i=1}^{N} m(h, w, I, X; \theta) \right] \right) \tag{20}$$

where W is a matrix that weights the different moments.[21]

In the labor supply model of continuous choice it is natural to build moment conditions using the first order condition for the optimal choice of hours. With the utility function in (13), the optimality condition cannot be used to solve out analytically for h, but one can still use GMM to estimate θ by using $E[\varepsilon^u] = 0$ with $\varepsilon^u = \log(w_i) - \psi_0 - \psi_x X_i + \eta \log(w_i h_i + I_i) - \gamma(\log(h_i)$. If in addition to X, wages and non-labor income are exogenous, a vector of instruments $Z_i = [1, w_i, I_i, X_i]$ can be used to magnify the moment conditions and just identify the parameters $(\psi_0, \psi_x, \eta, \gamma)$.[22]

In the case of the discrete choice model in 9 one can follow Avery et al. (1983) and

[20]it is also possible to use functions of the instruments $q(Z)$ to increase the number of moment conditions.

[21]The choice of of moments $m(\cdot)$ and moment-weight matrix W obviously affects the numerical estimates obtained in finite samples but do not affect the consistency of the estimator. We refer the reader to the more specialized discussion of GMM available in several econometrics textbooks, specially for details on how to best choose the matrix W and how to compute standard errors.

[22]To identify σ_u^2 one can also use the moment given by the variance $E[(\varepsilon^u)^2 - \sigma_u^2] = 0$.

represent discrete choice problem in a moment framework where $E[d_i - E[d_i|Z_i, \theta]] = 0$. Note that because d_i is binary, the moment condition is equivalent to $E[d_i - \Pr(d_i = 1|Z_i, \theta)] = 0$

3.5.2 Simulation-Based Estimation

As described in Stern (1997), simulation-based estimation methods are an increasingly popular alternative for estimation of structural models. In many cases, the classic methods described in the previous section become intractable and a simulation approach is often more convenient. For example, the choice probabilities in models with many choices are difficult to evaluate and simulators for them are instead very simple to implement. Similarly, whenever some choices or states are missing one might need to integrate over them. These integrals are often more easily handled via simulation. In this simulation-based approach, at each trial of the structural parameters, the model is solved and simulated. The simulated data or its moments are compared to the empirical data. The structural estimates are obtained when the data simulated from the model matches the empirical data. Lerman and Manski (1981), McFadden (1989) describe the simulation-based counterparts to classic approaches such as maximum likelihood and method of moments. An alternative approach is indirect inference as described in Gourieroux et al. (1993). Another driving force behind the trend towards the use of simulation-based estimation methods (e.g. Simulated Methods of Moments, Indirect Inference) is that they easily allow for combining multiple data sources as noted by Low and Meghir (2017). Re-weighting might be necessary to ensure that the different data sources represent the same population. As noted early on by McFadden (1989) and Stern (1992) smoothing the simulators used in simulation-based estimation might be important to exploit gradient-based optimization algorithms. Bruins et al. (2018) and Sauer and Taber (2021) provide alternative smoothing approaches in the context of indirect inference. A novel, promising approach to simulation-based estimation has been recently advanced by Kaji et al. (2020), integrating classic model-based simulation with techniques from machine learning that are used to discriminate whether an observation is real empirical data or simulated from the model. The structural parameter estimates in this approach are those that render the machine learning discriminator unable to distinguish simulated from empirical data. Eisenhauer et al. (2015) caution against the use of simulated methods of moments and emphasize the need for dynamic moments for estimation of dynamic models

Berkovec and Stern (1991), French (2005), Dey and Flinn (2008) provide applications of simulated methods of moments, whereas Keane and Wolpin (1997) and Dey and Flinn (2005) include applications of simulated maximum likelihood. Examples of applications using Indirect Inference include van der Klaauw and Wolpin (2008), Tartari (2015), Adda and Dustmann (2021).

In the context of the labor supply model in (16) one can simply obtain simulated versions of moment-based or likelihood-based estimators by simply replacing the model-based probability that individual i chooses to work, $\Pr(d_i = 1|w_i, I_i, X_i; \theta)$ with a simulated version of it, $\hat{\Pr}(d_i = 1|w_i, I_i, X_i; \theta)$. Suppose we take, for each individual i, R draws of ε^u from its distribution. Then a crude frequency simulator $\hat{\Pr}(d_i = 1|w_i, I_i, X_i; \theta)$ is given by

$$\hat{\Pr}(d_i = 1|w_i, I_i, X_i; \theta) = \frac{1}{R}\sum_{r=1}^{R} \mathbb{1}\left[U(w_i + I_i, 1, X_i, \varepsilon^u_{i(r)}; \theta) > U(I_i, 0, X_i, \varepsilon^u_{i(r)}; \theta)\right] \tag{21}$$

where $\mathbb{1}[\cdot]$ is an indicator function that equals one when the statement or event in brackets is true and zero otherwise. Intuitively, the probability of working is replaced by the simulated fraction of times the individual would choose to work under a large, representative sample of the ε^u she might actually face. Stern (1997) provides additional detail on more efficient simulators and how to compute standard errors for simulation-based estimators.

Indirect Inference A a more recent practice is to rely on indirect inference estimators. In this approach one estimates a set of auxiliary models using the empirical data. These are just linear or non-linear statistical relationships between the observable variables and they are not attached any causal or structural interpretation.These auxiliary models typically use the model's endogenous variables as outcome variables and the model's state variables as explanatory variables. One can also here include among the auxiliary model, some difference-in-difference specifications or IV regressions that might provide a key source variation that one may want the model to match. Let the set of parameters from these auxiliary models estimated in the empirical data be denoted by π^{data}. An indirect inference estimator then proceeds as follows:

1. Guess θ.

2. Solve the structural model at that θ (i.e. find cutoffs like ε^{u*} in a discrete choice model or evaluate the first order conditions using θ in continuous choice models.) This tells us how the individual would behave for each ε^u she might face.

3. Take R simulated draws from the distribution of ε^u for each individual.[23]

4. Use the optimal behavior prescribed by (2) above to simulate endogenous choices that each individual would make given the different $\varepsilon^{u*}_{(r)}$.

[23]Ideally the underlying draws for each individual remain the same but their distribution gets updated when the distributional parameters get updated. For example, one can draw uniforms once and then convert them to normal draws with the mean and variance that are being proposed as the estimation unfolds.

5. Collect the simulated data for all individuals into a simulated dataset with R-by-N simulated observations.

6. Estimate the same set of statistical descriptive models on the simulated data and obtain $\pi(\theta)$.

7. Compare the vector π^{data} with $\pi(\theta)$ and if the difference is not very small go back to step (1) and guess a new θ.

In sum, the indirect inference estimator then solves:

$$\hat{\theta}_{II} = \arg\min_{\{\theta\}} \left(\left[\pi^{\text{data}} - \pi(\theta)\right]' W \left[\pi^{\text{data}} - \pi(\theta)\right] \right) \tag{22}$$

where W is a weight matrix.[24]

3.5.3 Unobserved Heterogeneity

These models often are estimated by allowing for unobserved heterogeneity. There are different approaches to allowing for unobserved heterogeneity. One popular approach is to allow for a discrete distribution of unobserved types and letting some of the structural parameters to vary by type. Heckman and Singer (1984) first proposed this method in the context of duration models but it has since been applied to structural models more generally. One then jointly estimates the probability of these types along with the type-specific structural parameters in the same estimation routine. Consider a dynamic extension of the discrete choice model in (9). To focus on the central issue, let's ignore the modeling of wages, but allow instead for K types of individuals sharing a permanent component μ_k^u in the distaste for work $\varepsilon_{i,t}^u = \mu_k^u(i) + \nu_{i,t}$ as discussed in Subsection (3.3). Observing panel data of length T from individuals one needs to take into account the dependence of choices over time introduced by the permanent unobserved types. Given that $\nu_{i,t}$ are independent not only across individuals, but also for a given individual over time, the joint probability of the observed history of choices for given individual, assuming that she is of a given type k, simplifies to the product over time of the type-k-specific individual choice probabilities. The type-specific individual likelihood function then becomes

$$L_i(\theta; k, Y_i, X_i) = \prod_{t=1}^{T_i} \Pr(d_{it}|k, w_{it}, I_{it}, edu_i) \tag{23}$$

[24]For additional technical detail on indirect inference and associated standard errors, see Gourieroux et al. (1993).

and the likelihood contribution for a given individual just integrates out the type-specific likelihood contributions

$$L_i(\theta; Y_i, X_i) = \sum_k \Pr(k) L_i(\theta; k, Y_i, X_i) \tag{24}$$

The likelihood function for the whole sample is then given

$$L(\theta; Y_i, X_i) = \prod_{i=1}^{N} \sum_k \Pr(k) L_i(\theta; k, Y_i, X_i) \tag{25}$$

In the context of dynamic programming models, one might also let the transition probability for the state variables like the one in Equation 12 to depend on the same permanent unobserved types that preferences are allowed to vary by. This becomes quite computationally costly as one needs to re-solve dynamic programming models during the estimation routine also for each change in the parameters of Equation 12 even when the parameters of preferences remain the same. If Equation 12 does not depend on the same unobserved types, one can estimate its parameters in a first step and "plug them in" into the second step, simplifying the estimation procedure substantially. To deal with this problem Arcidiacono and Jones (2003) obtain computational savings by adapting the EM-algorithm and converting the estimation procedure into an iterative process in which simpler and faster estimation steps can be conducted in each iteration by breaking down the estimation in each iteration into separate steps for transition and preference parameters with an extra step that updates the probability that a person is of each type. A common procedure to select the number of types is to start with 2 or 3 and increase the number of types until the estimation criteria (e.g. the likelihood function) does not substantially improve. Kasahara and Shimotsu (2009) consider the non-parametric identification of these models.

An alternative approach is to let the structural parameters to be continuously distributed in the population. One then has a continuous joint distribution for the parameters that are allowed to be heterogeneous. A parametric form such as a multivariate normal is often used. One would then rely on simulation-based integration, to integrate out this "continuous" unobserved heterogeneity in the estimation routine. This typically requires many more solutions to ensure that those simulators provide good approximations and this is why the discrete type approach is more popular.

An alternative approach to allow for unobserved heterogeneity uses unobserved factors to connect various parts of the model and flexibly accommodate additional sources of correlation in behavior by allowing, for example, for the same unobserved factors to enter choice and outcome equations with different factor loadings thus providing a model-based approach to dealing with a typical form of endogeneity. Carneiro et al. (2003) and Aakvik et al. (2005) discuss identification of these models when the unobserved factors are continuously distributed.[25] Cameron and Heckman (1987) and Mroz (1999) consider the discrete factor case.

The discrete "unobserved factor" and discrete "types" approaches are also useful devices to deal with the initial conditions problem. This issue is often encountered when there is unobserved heterogeneity and behavior preceded the time window available in the sample. As pointed out early on by Heckman (1981a), in such situations the initial state is not exogenous in that, by the very logic of the model's dynamics, certain types will more likely to have certain initial conditions. A common solution is to let the unobserved type probabilities be a function of the observable initial conditions.

With the advent of parallel processing techniques, recent approaches that allow for even richer forms of unobserved heterogeneity have been proposed by Ackerberg (2009) and Fox et al. (2011). In these approaches, many solutions to dynamic programming problems are first computed under different values for the structural parameters (exploiting parallel processing, if available). One then simply re-weights the contributions of these solutions when integrating out the unobserved heterogeneity during the estimation routine avoiding new computationally costly solutions to the dynamic programming problem. Estimation proceeds much faster than with traditional algorithms because those algorithms need to resolve multiple dynamic programming problems each time new parameters for the distribution of unobserved heterogeneity are proposed in estimation. While these methods have been more often used in the context of empirical industrial organization applications, we see no reason why they couldn't be applied to research on the themes of this Handbook.

3.5.4 CCP Estimation

Dynamic structural models of discrete choice quickly become computationally intractable, when one allows for several state variables in the model. Building on the work of Hotz and Miller (1993) there has been a recent trend towards the use of structural estimation methods for dynamic models of discrete choice that avoid the solution to the dynamic programming problem by exploiting a representation of the value functions that only depend on conditional choice probabilities (CCPs). These CCPs can be estimated directly from the data and are kept fixed during the estimation routine. Hotz et al. (1994), Aguirregabiria and Mira (2002) and Bajari et al. (2007) provide extensions of this approach that improve its practical applicability. Altug and Miller (1998) extend the CCP approach to allow for aggregate

[25] See also Heckman (1981b) and Aakvik et al. (1999)

shocks and Gayle et al. (2018) extend the CCP approach to dynastic settings. While an early concern with the approach of Hotz and Miller was that it could not handle unobserved types as in Subsection 3.5.3, Arcidiacono and Miller (2011), Pantano and Zheng (2013) and Bonhomme et al. (2021) propose estimation methods that can be used to extend the Hotz and Miller (1993) approach to models with unobserved heterogeneity in structural parameters. Gayle (2018) extends the ideas in Arcidiacono and Miller (2011) to models with both discrete and continuous choices. Most of this line of work uses models where the unobservables are additively separable. Kristensen et al. (2015) extend the CCP approach to models with non-separable unobservables. While the CCP approach can provide computational savings of several orders of magnitude during the estimation, it remains the case that one must, once estimation concludes, actually solve the model at the estimated parameters in order to use it for, say, evaluating counterfactual policies. A thorough discussion of the CCP approach is beyond the scope of this review. We refer the reader to comprehensive surveys by Aguirregabiria and Mira (2010) and Arcidiacono and Ellickson (2011) for more details.

3.5.5 Measurement Error

As mentioned at the beginning of this section, an alternative approach to structural estimation is based on the premise that the observable variables in the model are sufficient to capture all that's relevant for the decision-maker to make choices. From this standpoint, the model is correct and any discrepancy between model predictions at the individual level and actual microdata must be due to measurement error in the microdata. In this approach, an alternative way is proposed to rationalize why individuals with the same observables have different wages and even those with the same observables and wages end up making different work hours choices. Instead of relying on structural unobserved heterogeneity $(\varepsilon^u, \varepsilon^w)$ one argues that, say, hours and wages are measured with error. The measurement errors can rationalize some departures between the data and the model without the need to augment the stochastic structure of the model that the individual is assumed to observe. One must however be rather disciplined when allowing for measurement error and in contrast to structural features, not let the measurement error be too flexibly specified because no matter how bad a model could be, it could always be reconciled with the data if allowing for sufficiently flexible measurement errors. The typical approach is to assume that the observed values of y are the true values y^* contaminated with measurement error, ε^y as follows

$$y = y^* + \varepsilon^y \tag{26}$$

One can then postulate a distribution for the measurement error ε^y and estimate its parameters along with the other model parameters, using for example, a maximum likelihood approach. In this approach, the maximum likelihood estimator tries to jointly choose structural parameters that capture the main patterns of behavior and parameters for the distribution of the measurement error that could reconcile the observed deviations between model predictions and the data at the individual level. Early examples of the use of measurement error in estimation include Wolpin (1987), Eckstein and Wolpin (1989a) and Stern (1989), all of which allow for measurement error in wages. Less common is to allow for measurement error in discrete variables. For example Keane and Wolpin (2001) allows for classification errors in the measurement of the model's discrete choices as well as measurement error in continuous states such as wages and assets.[26]

3.6 Validating Structural Models

Once the structural model has been estimated, it is customary to asses how well it fits the data that was used to estimate it. It is common to report tables and figures showcasing how well the predictions of the model match choices and outcomes under the baseline or status-quo environment. This contrasts to models typically used in macroeconomics that rely more heavily on aggregate data. A feature of the structural micro-econometric approach is that this model fit evaluation process can be more thorough and demanding. One can examine not only whether the model fits means and standard deviations for choice variables but also examine covariances. Any moment or statistic that one can construct from the microdata can be in principle compared to a similar statistic computed from the model, either analytically or by simulation. A more exacting way of validating the model is to see how it matches behavior out of sample. This is what is called out-of-sample validation. To conduct this type of model validation one must have access to data that is "held-out" and is not used for estimation. Unlike the validation approach used in machine learning that sets aside a hold out representative sample for validation purposes, here the idea is to reserve a sample that was exposed to different incentives. For example, one could use a sample that faced a complex set of tax and transfers and see if the simple model above estimated on a sample of individuals, that came from the same population, but that were not taxed and did not have access to any transfers, matches the behaviors of individuals subject to taxes and transfers when these are simulated within that model. The idea is that finding that the model does indeed match the behavior under a different environment would convey a high degree of credibility to the estimated model. We defer further discussion on the issue of external or

[26] See also Imai and Keane (2004).

"out-of-sample" validation to Subsections 5.2 and 5.3 where we take it up again, in the context of discussing the integration of structural and experimental and quasi-experimental approaches.

4 Advantages and Disadvantages of Structural Models

In this section we discuss advantages and disadvantages of structural models. We also discuss, separately, the more debatable issue of external validity, which under some circumstances, can be seen as another advantage of the structural approach.

4.1 Advantages of Structural Models

Structural models are often difficult to estimate so it is important to have a clear understanding of their advantages over simpler empirical approaches that may offer more transparent solutions at lower programming and estimation time costs. The three clear advantages of structural models are a) the ability to simulate behavior under new environments or policies that have never been experienced by the population under study, b) assess the importance of various mechanisms and c) evaluate the welfare implications of alternative policies or changes in the environment that individuals face. We discuss each in turn.

1. **Ex-Ante Evaluation of New Environments** In some cases, the answer to the research question of interest is the set of structural parameters themselves. For example, one might be interested in certain feature of the utility function or in the productivity of certain input in a production function. In those cases, the estimated parameter directly provides the answer to the research question. Most often, however, the question of interest is how the decision makers would behave under alternative environments, including, as special case, policy changes. These policy or environmental changes can be easily evaluated using the structural model once the primitives have been estimated. In many cases the evaluation of this policy changes involves simulation of unobservables that enter the agent's decision-making process but that are not observed by the econometrician. The focus in this counterfactual experiments is often on how choices and outcomes Y differ in this alternative environment relative to the baseline or status-quo. The baseline environment is the environment that decision-makers faced when the data used to estimate the model was generated. The model we have been discussing applies to a hypothetical population that did not have taxes nor transfers. Once we have recovered the structure we can use that model to evaluate how the labor supply of this population would change if exposed to a complex labor income tax and transfer

system. All we need is to set up a new budget constraint that incorporates the taxes τ and transfers of interest B:

$$C = (1 - \tau(w_i, h_i))w_i h_i + B(w_i, h_i) + I_i \tag{27}$$

2. **Mechanisms** Structural models have the ability to evaluate the importance of various channels or mechanisms. For example, one can augment the labor supply model above with child development where the individual cares about the level of development of her child, Q. One can adjoin a technology of development where Q depends on how much time t^Q the mother spends with the child and how many development enhancing goods c^Q she purchases in the market. The model is then modified as follows: the new utility function depends on child development $U(c_i, l_i, X_i, Q_i, \varepsilon_i^u)$ so a new preference parameter must be considered, capturing how the mother trades off her consumption and leisure and the development of her child. The time constraint is modified so that $T = h_i + l_i + t_i^Q$. A production function for Q is added to the structure of model, $Q_i = q(t_i^Q, c_i^Q)$. A researcher can estimate this model with access to additional data on $\{Q_i, t_i^Q, c_i^Q\}$. Then one cannot only simulate the effects of a welfare benefit on child development, but one can also distinguish how much of the total effect comes from the additional developmentally enhancing goods c_i^Q that the mother chooses to purchase, and how much comes from the increase in time that the transfer allows her to spend with her child, reducing her labor supply.

3. **Quantifying Welfare Effects** Structural models often have the ability to provide a monetary measure of the value that individuals attach to certain changes in the environment. Let $V(w, I, X, \varepsilon^u)$ be the indirect utility function associated with the simple labor supply model in (2). $V(w, I, X, \varepsilon^u)$ can be obtained by using the optimal hours of work to recover optimal leisure, l^* and consumption c^* and plugging them into the utility function $U(\cdot)$. Once the structural model is estimated, it is easy to compute V using the estimated direct utility function U. Let's further index the "status quo" or baseline environment that generated the data used in estimation by e^0. One can derive the monetary (positive or negative) willingness to pay ($\texttt{WTP}_i$) that each individual has for a new environment e^{new} by simply finding the value $\texttt{WTP}_i$ that solves

$$V(w_i, I_i, X_i, \varepsilon_i^u; \boldsymbol{e^0}) = V(w, I - \texttt{WTP}_{\boldsymbol{i}}, X, \varepsilon^u; \boldsymbol{e}^{\text{new}}) \tag{28}$$

4.2 External validity

We devote a separate section to the issue of external validity because there is some debate on whether it can be considered an advantage of the structural approach. We discuss why the claims of external validity often credited to structural models must be qualified.

Structural Choice Models. While it was at some point conventional wisdom to argue that structural models have the additional advantage of providing external validity, this is no longer so clear in a world that allows for unobserved heterogeneity in preferences. Consider for example, the labor supply model above with $K = 2$. The are two types. For example, there could be a lazy type ($k = 1$) with a high value of μ_k and an industrious type ($k = 2$) with a low value of μ_k. The type of each individual is unobserved to the econometrician. From the econometrician's perspective each individual is of the lazy type with $\Pr(k = 1)$ and of the industrious type with probability $\Pr(k = 2) = 1 - \Pr(k = 1)$. Once preferences are heterogeneous, the structural model recovers the distribution of preferences for the population from which the sample was drawn. Suppose for the sake of the argument that we estimate the lazy and industrious types to be evenly distributed in the population from where the sample to estimate the structural model was drawn. There is no guarantee that that is the right distribution of preferences in another population. In fact, there is mounting evidence on how different preferences are across populations. Recent evidence by Falk et al. (2018) document striking patterns of heterogeneity across, and especially within, countries in measures closely associated to structural parameters such discount factors and levels of risk aversion, altruism and social preferences. Our estimated model, even if well-identified in the estimating population may do a poor job at forecasting the response to policy in another population that has an unknown, different distribution of lazy/industrious types. One possible way around this is to assume that differences across populations in their unobserved heterogeneity distributions can be captured by observable variables. In that case one can parameterize the distribution of unobserved types in terms of exogenous observables $\Pr(k|X)$. Then with access to data on X in the new location, one can derive what the prevalence of unobserved types is there, and thus construct external validity for the estimated structural model. In our view, then, structural models do provide some external validity but in a more restricted sense than traditionally claimed. When unobserved heterogeneity in structural parameters is very important, external validity might be difficult to claim as an advantage of structural models. These models can predict behavior under a new counterfactual environment in the same population but might be unable to predict behavior under the same environment (let alone a different one) in alternative populations whose unobserved heterogeneity distributions are unknown. Note that this is a problem that becomes more visible by estimating a more realistic model that allows for unobserved heterogeneity in the population in which the model is estimated. If one assumes instead that every person in the estimating population has the same structural parameter (or that all differences in those parameters can be accounted by observable variables) then one might incorrectly assume that individuals in other populations might also have the same parameter and the model is externally valid. Not allowing for unobserved heterogeneity in structural features essentially assumes away the problem. It does not fix it.

Similarly, it is not so clear whether one can extrapolate to other time periods, out of the sample window, even for the same population used in the estimation of the structural model. The traditional view is that structural models could in principle do that as well. However, any out of sample time effects cannot be identified by any method, whether structural or not. For example, suppose the unobservable distaste for work may be subject to aggregate shocks at different points in time, capturing underlying aggregate changes in the opportunities to enjoy leisure $\varepsilon_{it}^{u} = \mu_t + \nu_{it}^{u}$. Then, the estimated structural model estimated at time t will not be able to correctly predict behavior at time t' whenever $\mu_t \neq \mu_t'$. In fact, this is why ex-post evaluation methods that control for time effects such as difference-in-differences approaches are so popular, even when they do not enjoy the advantages discussed in the previous section. Our view is that even when the identified structural model is still right, it may not predict behavior appropriately when taken to another point in time (in the future or in the past) that features an unknown and distinct time effect.

Structural Models of Treatment Effects An important set of structural models adjoins an outcome equation to the structure that characterizes how an individual decides to participate in a treatment that has heterogeneous effects on an outcome. These structural models of treatment choice which model the potential outcomes under each possible treatment can be used to construct the full marginal treatment effect (MTE) profile. As described in Heckman and Vytlacil (1999, 2001, 2005), once one recovers the MTE, one can then estimate *any* treatment effect in the population under study. It is often argued that the LATE parameter of Imbens and Angrist (1994), which is identified by reference to a given binary instrument might not be externally valid. A common concern is that this parameter may not provide a relevant estimate for the average treatment effect in the whole population, the average treatment effect on the treated or the average treatment effect among those who might be induced to take treatment by a policy of interest. We share this concern. An argument is then often made that, since a structural model of treatment effects can in principle recover the whole MTE profile even when relying on a binary instrument, such a model can, as the argument goes, obtain *any* treatment effect in the same population at the same point in time, not just a LATE for the Z-compliers. One might then be tempted to argue that the structural approach has more external validity relative to a clean Wald estimate of LATE that imposes the minimal structure. Yet, it is important to recognize that, as emphasized by Kline and Walters (2019), this ability to generalize beyond the complier sub-population often relies on extrapolations that might be heavily dependent on functional forms assumptions whose validity might be questionable. So even this more limited sense of external validity that captures ability to generalize to other sub-populations (but within the population under study and at the same time and in the same place) other than the Z-compliers comes at the price of having to maintain functional form assumptions. So while it might be true that a non-structural estimate of LATE might not carry external validity, it is also true that the more structural models of treatment choice and treatment effects attain that wider reach beyond the complier sub-population by relying on potentially questionable functional form or distributional assumptions.

In sum, in our view, a structural model is best at predicting what would had happened at the *same* time, in the *same* place, had the *same* population faced an alternative environment. Furthermore, accomplishing this goal may often come at the price of having to maintain functional form assumptions that are necessary for extrapolation. This is an issue that we return to when discussing out-of-sample validation in Section 5.2.[27]

4.3 Disadvantages of Structural Models

Discussions about the relative value of structural models are often lopsided with advocates championing its use and detractors emphasizing their disadvantages. Our goal here is to provide a balanced, level-headed review of its pros and cons. Having described the unique advantages of the structural approach it is now time to be explicit about its disadvantages. This not only serves to provide a more neutral review but also dovetails nicely into our next section where we describe a recent trend that aims to address some of these disadvantages. Some of the ideas in this section draw from Angrist and Pischke (2010), who provide a blunt criticism of the structural approach.

Identification It is argued that in these models it is difficult to prove formal econometric identification in the sense of Matzkin (2007, 2013). As argued in Section 3.4, structural studies rarely offer a formal proof of identification. Even putting that aside, another concern is that empirical identification might be achieved by exclusion restrictions that are not as rigorously vetted. Papers that are built around estimation of a LATE with a convincing instrument or fixed-effects models often devote quite a bit of time and space to justify the

[27]See Banerjee and Duflo (2009), Imbens (2010), Bo and Galiani (2020) and List (2020) who tackle the issue of external validity from complementary perspectives.

validity of the identification adopted. In stark contrast to this focus, the exclusion restrictions in some of the early structural literature do not appear to have been as well though-out, and were often relegated to footnotes or data appendixes when explicitly reported at all. Coming up with an exclusion restriction that was key for identification was more of a formality than a substantive, critical empirical problem that defined the whole empirical strategy. For example, it is possible to show that in the simple binary choice model of labor supply, it is necessary to have a variable affecting wages that does not affect taste for leisure. Arguing that education could play such a role, affecting wages but not distaste for work is something that perhaps the early structural literature may not have had issues with, but those coming from a more recent non-structural perspective are now less prepared to accept. As we argue below though, this is not a fundamental problem with the structural approach. It is just that, nowadays, the bar is set much higher for what is considered a valid exclusion restriction. Indeed, we discuss in the next section how the modern structural approach can adopt this higher standard by integrating sources of experimental or quasi-experimental variation for identification.

Functional Form. Structural models are correctly seen as heavily-reliant on parametric functional forms. While there is some theoretical work on non-parametric structural models, most of the applied work is heavily parametric. Indeed, even in the very few cases in which full non-parametric identification of the structural model is provided, researchers often go on to estimate the model under much more tightly parameterized functional forms for utility functions, production functions, and the stochastic structure, without necessarily testing the validity of the parametric restrictions against the more flexible non-parametric structure that can be in principle identified under ideal data conditions. Even if one had access to ideal data conditions and could estimate the model non-parametrically, a parameterization will often have to be made anyways whenever the counterfactual of interest is somewhat outside the support of the data. A non-parametric approach is silent about what the structure looks like beyond the support of the data and is thus unable to help in extrapolation exercises. While the need to impose functional form to extrapolate seems an inherent feature of the structural approach, we discuss in the next section how one can use experimental or quasi-experimental sources of variation to externally validate the model and help select appropriate functional forms in some cases.

Computational Complexity Finally, while this is starting to change somewhat, it is well-known that the necessary programming for estimating structural models must often be coded from scratch, without the possibility of relying on canned software packages. This is often very time consuming because: a) programming b) debugging c) running the code and d) making changes to the model to try new specifications all may take a long time. All of this time-consuming activities take away from the researchers' available time to focus on the economics of the problem and to search for more exogenous sources of variation. It also prevents extensive sensitivity analysis and skeptics are often left to wonder, and rightly so, how robust the results might be to even small changes in many of the model's details. This is something that seems quite inherent to the approach, although there have been attempts to code general purpose packages that can be used for different applications. However, these have not been adopted widely as the details of these models often mean that a generic code will not be sufficient to capture the idiosyncrasies of a particular application. Nevertheless, similar in spirit to the CCP estimation approach discussed in Section 3.5.4 there is continued methodological work to develop methods that allow the researcher easier ways to estimate structural models. Eberwein and Ham (2008) illustrate the substantial computational savings that can be achieved by using analytic instead of numerical derivatives when estimating dynamic structural models of discrete choice via maximum likelihood. They also show how analytic derivatives can help in debugging code and easily spotting programming errors.

5 The Integration of Design-based and Structural Approaches

The last few years of the twentieth century and the first decade of the twentieth-first century witnessed increased specialization of empirical research in applied microeconomics. Most researchers seemed to take two quite different routes to conduct empirical work. A popular approach, following the credibility revolution described in Angrist and Pischke (2010), focused on finding a convincing identification strategy and building the research around it. It was increasingly common to see instrumental variables, regression discontinuity designs and difference-in-differences approaches, as alternative tools within this broad movement that did not emphasize the need for specification, let alone estimation, of a fully structural model, but rather, placed the research design front and center. Other sharing this focus on "clean identifiaction" went on to design experiment themselves to generate the needed variation, particularly in the field of development economics.[28] An alternative, and certainly less common strategy, was to follow the approach we describe in this chapter, opting for spending most of the research effort formulating and estimating fully specified structural models where the source of exogenous variation was, in relative terms, a less central concern.

Against this backdrop of increasing methodological specialization and polarization, the

[28]Perhaps due to larger costs of conducting field experiments in developed countries, this has been somewhat less common in labor economics and other applied microeconomic fields. But see List and Rasul (2011) for a survey of experiments (including laboratory ones) in labor economics.

last 10 years have shown some signs that this trend might be reversing, with increasing attempts to integrate the two paradigms both in research and teaching. In this section we provide a brief review of this ongoing trend that seeks to integrate structural models with experimental or quasi-experimental sources of exogenous variation, which are the hallmark feature of non-structural or "design-based" empirical strategies. Rather than cataloguing every paper that has integrated these approaches in one way or another, we believe it is more useful to try to provide a road-map to the distinct ways in which this integration is taking place. We believe this integration can be fruitful in addressing some of the disadvantages associated with structural models discussed in the previous section and reinvigorate the structural approach. In assessing this recent literature, there seem to be currently two schools of thought developing on how this integration can best be implemented.

On one end some argue that these clean sources of variation should be used in the estimation of structural models. One possibility here is to use the experimental variation to identify a "stigma" parameter that can only be identified with the treatment-control contrast given by the experimental variation. These are parameters that are not relevant in the control group because they are only part of the structure when the individual is exposed to the treatment. NOte that these parameters are not necessarily negative. The "stigma" label follows from Moffitt (1983), who introduced the idea of individuals suffering a utility cost from participating in a welfare program. An alternative use of the experimental variation, when these types of "stigma" effects are thought not to be important, is to use the experimental variation as an instrument to relax a debatable exclusion restriction. For example, as discussed in Section 4.3, assuming that wage is affected by education but distaste for work is not would be a debatable exclusion restriction nowadays. If one has access to a Randomized Control Trial (RCT) that assign wage subsidies, one can rely on it for identification and let education affect both wages and distaste for work.

An alternative line of thought argues instead that it might be best to hold out such variation from estimation, reserving it instead for external, out-of-sample validation of the structural model. In this section we complement a recent review by Todd and Wolpin (2020) that focuses on this "out-of-sample" validation perspective by emphasizing why, whenever possible, it might be best to do *both*: use some of the available variation to estimate the model and some to validate it, as in Galiani et al. (2015). Further, we discuss a third way in which this combination is fruitful: facilitating the unpacking of bundled features of treatment.

5.1 Using Experimental and Quasi-Experimental Sources of Variation to Estimate Structural Models

In this subsection we discuss how design-based approaches are being leveraged for identification of structural models. There is a hope that by using this type of variation the identification of θ might be more convincing or credible, in the same sense that an ATE estimated using an RCT with perfect compliance or a LATE with valid instruments are thought to be convincing.

Imbens (2010) provides a compelling call for the use of experimental variation in estimation of structural models. He argues that even though structural models have many parameters and one cannot hope that a simple two-arm RCT will provide distinct variation to identify them, it might at least help identify a combination of them imposing some discipline in the identification of the structure. Similarly, Heckman (2010) welcomes the use of experimental variation for identification of the marginal treatment effect (MTE) profile. While not a fully structural approach, in the sense of this chapter, estimating the MTE may in many cases allow for the analysis of the impact of alternative policies as long as the policy variation in the experiment or any additional instruments used to estimate the MTE is not too different from the policy question of interest. In other words, one can re-weight the MTE to obtain a policy-relevant treatment effect when the policy of interest can be re-casted in terms of alternative configurations within the existing instrument-induced variation.

One early example of work that uses randomized experimental variation to estimate a structural model include Burtless and Hausman (1978) who exploit data from the Negative Income Tax experiments to estimate a structural model of labor supply.[29] More recently, Imbens et al. (2001) use data from lottery winners to estimate a dynamic model of labor supply whereas Ferrall (2012) use experimental variation from Canada's Self-Sufficiency Project to estimate a model of welfare participation, Attanasio et al. (2012) use RCT data from Mexico's PROGRESA conditional cash transfer program to estimate a model of child school attendance and Galiani et al. (2015) use data from the control group and the restricted experimental group in the Moving to Opportunity experiment to estimate a model of neighborhood choice. Chaparro et al. (2020) uses randomized variation from the IHDP program to estimate a structural preference parameter characterizing maternal parenting exhaustion in a model of early childhood cognitive development.

[29] However, Burtless and Hausman (1978) do not proceed by specifying and then estimating a direct utility function but rather start with a conventional labor supply function for hours of work and back out the implied indirect utility function using Roy's identity. Unfortunately, the associated direct utility function cannot be recovered from the indirect utility they obtain and this places some limitations on the type of analyses that one can pursue. Moffitt (1979) provide further analyses of the Negative Income Tax experiments.

Others have estimated structural models by exploiting difference-in-differences type of variation in estimation to help identify the structural parameters in population economics, labor economics and human resources. For example, Voena (2015) uses cross-state variation over time in changes to laws regulating grounds for divorce and property division upon divorce to estimate a dynamic model of married couples decisions about the wife's labor supply and the couple's savings. Similarly, Blundell et al. (2016) exploit policy variation across cohorts in the taxes and welfare benefits they faced to estimate a rich model of the behavior of women in the U.K. They use the model to analyze welfare policy. Also, in the same spirit, d'Haultfoeuille and Février (2020) exploit an exogenous compensation contract change in the French government's statistical agency to estimate a model of asymmetric information and derive the optimal compensation contract.

There have been calls too for the integration of the the LATE framework into structural models. IV and in particular the LATE framework developed in Imbens and Angrist (1994) has been widely popular among empirical researches due to its simplicity and the small number of stated assumptions.[30] In this regard, and following an earlier desideratum by Angrist and Pischke (2010) on the need for estimates derived from structural models to "line up" with those obtained under weaker assumptions, Kline and Walters (2019) argue for using instrument variation in the estimation of structural models. They go on to suggest that "model-based" or structural LATEs could and should be routinely derived from a structural model and compared to "unrestricted" LATEs to give more confidence to the structural model when the estimates match. It is important to bear in mind that for this "quality control" to make sense, one must first verify that the proposed structural model does satisfy the monotonicity condition that is necessary in the LATE framework.[31] In a handful of cases, estimating a structural parameter might be quite simple. For example, Blundell et al. (1998) use the so-called "instrumented difference-in-differences (DDIV)" approach which uses difference in difference variation as an instrument and show that the key structural parameter that governs the labor supply elasticity can be estimated directly by DDIV.

[30] However many have pointed out disadvantages of IV such as failure of monotonicity, violation of exclusion restrictions, relevancy of complier sub-population Heckman (1997), Rosenzweig and Wolpin (2000), Keane (2010a), Heckman (2010),Heckman and Urzua (2010), Wolpin (2013). For a more optimistic assessment of this framework see Imbens (2010) and Angrist and Pischke (2010).

[31] Recent work by Mogstad et al. (2020) brings attention to some problems that arise with the monotonicity condition in multi-instrument settings.

5.2 Use of Experimental or Quasi-Experimental variation for Validation of Structural Models

An alternative approach to pursue the integration of structural and designed-based approaches is to use the clean variation for validation rather than estimation of the structural model. The idea of using RCTs as benchmarks for the evaluation of non-experimental estimators goes back to Lalonde (1986), who evaluated the performance of different estimators of treatment effects. Even earlier, work by McFadden (1977) and Wise (1985) used RCTs to evaluate non-experimental models of transportation mode and housing demand, respectively. This approach is perhaps best illustrated in the modern structural literature by the work of Todd and Wolpin (2006), and further formalized in Schorfheide and Wolpin (2012, 2016). In this approach the experimental variation is not used in estimation. Instead, it is reserved or "held out" to perform an out-of-sample validation of the model. The model is estimated only on the control group or, sometimes, on the treatment group and validated with the other group. The idea is that if the model can match the behavior of a sample that was set aside and that was exposed to different incentives, one gains more confidence in the ability of the structural model to predict behavior under other counterfactual environments. Todd and Wolpin (2020) provide a thorough review of recent research that adopts this approach.

In the same way that, as discussed in Section 5.1, different types of exogenous variation featured in design-based research can be embedded into structural model for identification purposes, they can also be used for validation. We focus here on the use of RCT for validation. Keane and Wolpin (2007) and Galiani et al. (2015) argue that RCTs provide the first best validation strategy because the unobserved heterogeneity distribution might be different in the validation group if that group has not been randomized.[32] Similarly, even when validating within the same "population", but at a different point in time that features different incentives, it is difficult for the structural model to account in an unrestricted way for time effects (e.g. anything that changes in the environment) that could be present. Keane and Moffitt (1998) estimate a model of labor supply and welfare program participation in the mid-1990s and validate it with an earlier cohort's behavior in the mid-1980s that faced different incentives. Again, they are relatively successful in that endeavor but nothing would prevent unaccounted cohort effects to hamper the predictions of their model. So failure to accurately predict behavior cannot be taken as evidence against the model's validity. In sum, only in an RCT one can ensure that the validation group that is held-out has the same

[32]For example, Choi (2018) estimates a simple labor supply model in one location and attempts and fails to validate it in another location. While Choi (2018) presents this as evidence of shortcomings of that simple structural model, it is also possible that failure to fit behavior in the validation group could just reflect a different distribution of unobserved heterogeneity in that group, rather than a mis-specified or invalid model.

distribution of preferences and is exposed to the same time effects as the group that is used for estimation.

In addition to the influential work of Todd and Wolpin (2006), subsequent applications include Misra and Nair (2011) who modeled salesmen effort and sales within a firm. Using their estimated model Misra and Nair (2011) provided recommendations to the firm about how to best structure their compensation contracts. Interestingly, the firm went on to implement their suggestion so they were able to validate their model by comparing model-predicted responses to the recommended contract with the actual responses observed when that contract was actually implemented by the firm. Duflo et al. (2012) used data from and RCT in India that introduced attendance incentives and monitoring to reduce teacher absenteeism. In most specifications they only use data from the treatment group for estimation, reserving the control group for model validation. More recently, Galiani et al. (2015) used data from the unrestricted (Section 8) treatment group in the MTO experiment to validate their model of neighborhood choice and Lise et al. (2015) calibrate a search and matching model to data from the control group in Canada's Self-Sufficiency Project and show that the model is able to match the treatment behavior of the treatment group when the same set of incentives that this group was exposed to are simulated within the model.

A more epistemological open question in this out-of-sample validation approach remains whether researchers should or should not have access to the validation group data while conducting research. A more stringent out-of-sample validity test is indeed given by one in which the researchers can never see the validation data so they are unable to go back and adjust the model specification in the hopes of improving the fit out of sample. Being able to access the data set aside for validation and modify the model upon iteratively until it fits amounts to actually using the validation data in the estimation in some sense. Arcidiacono et al. (2021) deal with this problem by splitting their research in two parts. They first estimate the model without seeing the validation data and pre-commit to this model. After this, they plan on accessing the validation data and write a second paper which will explore how well the model fits in the holdout sample.

Putting aside for the moment the question of whether this embargo of validation data should or should not be enforced, it is important to wonder whether it would be actually feasible. While in policy or industry contexts it is in principle possible to enforce such an embargo, it is more difficult to think about its feasibility in academic or scientific settings where the data should be available for replication.[33] Otherwise if a paper that does not fit out of sample is published, it would need to provide documentation of that failure, which others can then use to learn and improve subsequent models, again undermining the original pure out-of-sample validation intent for that data.

While this stringent "embargo" approach to external validation is a clever idea whose rationale has been formalized in Schorfheide and Wolpin (2012, 2016) as a device to prevent researcher data mining, we believe that the necessary institutions that would enforce the embargo of validation data might be difficult to implement in practice. There is also the question of how much information such institutions holding the validation data would disclose even if they are successfully established. If they only disclose that a model fails to fit but do not provide further details, scientific progress will be slower. The same researchers or other teams would only know that the proposed model was inadequate, but nothing else. They will not know in which dimension(s) the model fails to match, let alone whether it under- or overestimates- the validation sample's behavior in those dimensions.

So there is a tradeoff between the pace at which scientific progress will be allowed to unfold and the preservation of the validation data to maximize the purity of an out of sample test of a model's validity. It can evolve slowly if completely unaided by disclosure of validation data in a fully enforced embargo regime. But it will evolve nonetheless, because even the announcement that a proposed model failed to match out of sample would provide some minimal information about the validation data that can provide leads for how to formulate better models. But if many models with different predictions are rejected eventually researchers can trace out what the validation data looks like. So given that this scientific evolution will take place eventually and converge to a model that fits out of sample, one can in the other extreme do without the embargo and grant the initial research team immediate access to the validation data. They can then find the right model that fits out-of-sample themselves, subject to also fit in-sample.

We agree on the basic premise that there is much value in holding out a sample for blind validation. But there are institutional implementation issues to consider and the fact that even when those might be overcome, the embargo regime only slows down but cannot prevent the eventual discovery of the validation data. So we feel that it might be best for researchers to have access to these validation data and see how well the estimated model fits the holdout

[33] To implement a truly blind out-of-sample validation approach one might envision the creation of institutions within academia or within journals that would have exclusive access to these embargoed validation samples. These third parties would construct summary statistics with code provided by researchers, who would not be allowed to further modify the model at that stage. If the third party determines that the model fails to match out of sample, the paper would be rejected. Further, to prevent knowledge about the validation data to leak out, the authors of the paper would not be shown the validation sample statistics that document the failure of their model.

sample during the course of their research. If the initial model specification does not fit the out-of-sample data, they should be allowed to go back and refine the model. Model building in this context is thus an iterative process by which the researcher goes back and forth between model specification, estimation, model fit assessment both in-sample and out-of-sample, and then circle back to model specification, toggling between models with different functional forms or alternative non-nested models that may fit equally well in-sample but differ in their performance out of sample. One would then repeat these steps until the whole process converges to a model that fits well both, in sample and out-of sample.[34]. This approach thus alleviates some of the concerns about functional form discussed in Section 4.3. The confidence in the chosen functional form increases when, using this out-of-sample validation approach, one can show that it extrapolates correctly out of the support of the data.

In this out-of-sample validation approach without embargo, the holdout sample is actually used in an iterative process to select the right functional forms for competing non-nested models and which of these competing models is the most appropriate. But, conditional on the selected model, the experimental variation that this validation sample would offer is not necessary to identify the structural parameters. An important area for future methodological research is then the formalization of this iterative process and the understanding of its statistical properties, including the effects of data snooping on computation of standard errors, etc.

5.3 Use of Experimental or Quasi-Experimental variation for *both* Estimation and Validation of Structural Models

The last two subsections emphasized how structural models can benefit from integration with design-based experimental or quasi-experimental approaches for *either* identification or validation. Even when faced with the same RCT some researchers have opted for using the variation for identification whereas others opted instead to use it for validation. Take for example Canada's Self-Sufficiency Project. While Ferrall (2012) used the data for identification, Lise et al. (2015) opted to reserve it for validation.[35] Similarly, when using the *Progresa* RCT in Mexico, Attanasio et al. (2012) decided to use the experimental variation

[34]This is mechanically similar to the way that machine learning algorithms iterate between training and validation during the cross-validation process to fine-tune hyper-parameters or select learning algorithms. It is fundamentally different though, because the sample used for validation in a structural model faces different incentives.

[35]Card and Hyslop (2005) also anlayze the SSP experiment with a rich dynamic econometric model that allows for state-dependence and unobserved heterogeneity, but that differs from the more fully structural approach adopted in Ferrall (2012).

for identification whereas Todd and Wolpin (2006) opted instead to use it for validation.[36] It is difficult to provide a one-size-fits all answer to the question of whether experimental variation from an RCT should be used for identification or validation of a structural model. This should be judged on a case-by-case by comparing the marginal return in each of these two alternative uses.

Whenever possible though, it might be best to do both as in Galiani et al. (2015), where we take advantage of a 3-arm experiment with one control group and two treatment groups. The availability of more than one treatment group afforded us the opportunity to use some of the experimental variation for identification, as in Section 5.1, and some for validation, as in Section 5.2. In Galiani et al. (2015) we use data from the Moving to Opportunity experiment to estimate a model of neighborhood choice, and use the estimated model for ex-ante evaluation of alternative housing assistance policies. In the MTO experiments public housing residents were randomized into either a control group (C), and two treatment groups (T1,T2). The first treatment group (T1), was given vouchers to rent apartments in the private rental market, subject to the constraint that the voucher could only be used in neighborhoods where the poverty rate was less than 10%. They were also given housing mobility counseling. A second treatment group (T2) was just given unrestricted vouchers that could be used anywhere. We estimated the model of neighborhood choice using data from the control group (C) and the first treatment group (T1) and reserved the other treatment group (T2) for out-of-sample validation.

Suppose one has access to a 3-arm experiment with a control group (C) and two treatment groups (T1,T2). Following our discussion so far, there are in principle three strategies:

1. **Validation-Only:** Here, like in Section 5.2, the researcher would use only group C for estimation and reserve both T1 and T2 for validation.

2. **Identification-Only:** Here the researcher would use all data (C,T1,T2) for identification and estimation, holding out none for validation

3. **Identification+Validation** Here the researcher would use two of the three groups (e.g. C+T1, C+T2 or even T1+T2) for identification and estimation and reserve the third group for validation as in Galiani et al. (2015).

We believe that the third approach should be the preferred one in most cases. The hold out sample can be used to guide model selection and choices about functional form and distributional assumptions whereas the experimental variation that is dedicated to estimation

[36] Wolpin (2013) and Todd and Wolpin (2020) discuss in some detail the differences between Attanasio et al. (2012)'s approach and their own approach in Todd and Wolpin (2006).

can be used to better identify the model, by either relaxing a debatable exclusion restriction or using the variation to identify a richer model that incorporates parameters that can only be identified by contrasting the control group with one of the treatment groups. We believe this approach more thoroughly addresses some of the concerns laid out in Section 4.3.

Note also that while our focus here has been on the use of multi-armed RCTs for both identification and validation of structural models, other forms of exogenous variation can be used in the same way. For example, one could set aside some of the available geographic policy variation to estimate a model while reserving the rest to validate it, as in Keane and Wolpin (2007), who estimate their model with data from different states and then validate it with data from Texas, an outlier in terms of limited welfare benefit generosity. They succeeded in this endeavor, but had they not been able to fit the behavior in Texas, one would not be able to tell whether the model is invalid or there is just something different about Texas. Taking this idea across time, one can also use policy variation from some years to estimate the model while reserving the remaining temporal policy variation for validation as in Agostinelli et al. (2020) and Bobba et al. (2021). Again, they succeed in their out-of-sample validation exercise but, had they not been able to, one would not be able to tell whether the model was not valid or there is something different happening during the time period selected for validation. This is why, especially when the researcher has access to the validation sample, the RCT approach provides a more conservative setting to engage in an iterative process of model formulation and assessment.

5.4 Disentangling Bundled Features of Treatment

In the last few sections we have emphasized how structural models can benefit from an integration with experimental and quasi-experimental sources of variation to enhance their credibility. But how can experiments benefit from integration with structural models? In a series of papers, James Heckman has long emphasized limitations of RCTs (see, e.g., Heckman (1992), Heckman and Smith (1995), Heckman (2020)). Structural models can help address some of those limitations as well even when the goal is simply to learn about the treatment effect. That is even when no new policy or welfare question is the goal of the analysis. Another way in which a structural model can help in interpreting experimental results is by disentangling the separate effects of bundled features of treatment. It is common to argue that when a given treatment in an RCT has several "bundled" features, each of which is expected to have its own effect on outcomes, one can only obtain their combined net effect. It is not possible, the argument goes, to unpack the separate contributions of each of the features and their interactions (if any). It is often argued that the only way to identify these separate effects is by having multiple arms in the experiment with each feature separately randomized. This type of argument was common in, for example, the welfare policy experiments in the U.S. before the welfare reform of 1996. Many of these experiments that were conducted leading up to the reform had bundled features of treatment in the sense that the treatment group was simultaneously exposed to changes in the tax rate of welfare benefits and the disregard amounts or mandated work-related requirements.[37] The conventional wisdom was that only the bundled effect could be estimated when a single treatment group is simultaneously exposed to multiple bundled treatments. However, because of its ability to identify separate mechanisms, a structural model can often be used to tease out the separate impacts of each bundled feature of treatment, even when the experiment has a single treatment and control group. In other words, consider an experiment with only two arms: control and treatment. Suppose the treatment was a combination of two features that are expected to affect outcomes, say A and B. In that case one can use the structural model to disentangle the separate contributions of A and B to the treatment effects.

As described in Section 5.3, Galiani et al. (2015) estimated the model with the control group and a single treatment group (T1) that was given: A) a voucher with location restrictions for its use and B) mobility counseling. A traditional experimental analysis would only permit to analyze the combined impact of these two features of treatment on the rate at which public housing residents moved out of the public housing projects using the voucher. However our estimated structural model can be used to unpack the quantitative importance of each mechanism and investigate how each separate feature affected take up, and answer how take up would have responded if treatment had consisted of only the location-restricted voucher, without any mobility counseling. Similarly, the second treatment group (T2) that we held out for validation, received an unrestricted voucher but no mobility counseling and ended up taking up the voucher at a much higher rate. While a traditional experimental analysis could only identify the combined (opposite direction) impacts of the location restrictions and mobility counseling on the differential take up rate between T1 and T2, we were able to use our structural model to tease out how important each feature was. We found both features where quantitatively important, although in the end the location restriction on the voucher ended up dominating in magnitude and explaining why, on net, the T1 group took up the voucher offer at a much lower rate than the T2 group.[38]

It might be puzzling that a two-arm experiment might be able to identify two effects. This is certainly not possible when the two features of treatment have distinct stigma or

[37]See Grogger and Karoly (2005) for a comprehensive description of these experiments.

[38]The finding on the importance of counseling in Galiani et al. (2015) has been validated in a subsequent experiment by Bergman et al. (2020).

similar effects that expand the structure. However, in many cases, when at most one of the features is new and cannot be predicted by any model but the other feature of treatment could in principle be simulated within a model estimated only with the control group, then it is possible to tease out the two effects. For example, in Galiani et al. (2015) one feature of treatment was channeled through the budget constraint by reducing the effective rents that households faced in low poverty neighborhoods, whereas the other feature, mobility counseling, operated as a reduction in the utility cost from moving.

6 The Future of Structural Models

What does the future look like for structural models in empirical microeconomic fields? We envision a bright future for the structural approach. But structural methods could gain a lot from adapting to the new times and absorbing insights from other areas. We briefly discuss a few areas where we feel that more integration could be useful.

- **Exogenous Variation.** As emphasized in Section 5, first and foremost, we believe that the most important next step in the development of the structural approach is the more comprehensive integration of experimental and quasi-experimental variation for both the estimation and validation of structural models within the same research project. This practice is certainly underway and is something that's becoming more noticeable, particularly when comparing research articles from the last 10 years with the early, seminal structural literature from the 1980s and 1990s that initially broke ground without putting as much emphasis neither on validation strategies nor on the credibility of the exogenous variation or the maintained exclusion restrictions.

- **Expectations Data.** We feel that there might be high returns to more integration with the literature on subjective expectations pioneered by Charles Manski.[39] Even when assuming rational expectations, having access to subjective expectation data amounts to having a larger sample and can help improve the precision of structural estimates as in the work of van der Klaauw and Wolpin (2008) and van der Klaauw (2012) or help identify unobserved heterogeneity as in Pantano and Zheng (2013).[40] Most importantly, it could help relax the assumption of rational expectations as in Wiswall and Zafar (2015).

[39] See for example Manski (2008) for an early summary of this literature.

[40] See also Pistaferri (2003) on the creative use of subjective expectations to estimate the Frisch elasticity of labor supply.

- **Behavioral Economics.** Most of the early structural literature was firmly grounded in the mainstream economics of the time which postulated stable, time-consistent preferences, rational expectations and no barriers to dynamic optimization. These assumptions, which were relatively uncontroversial at the time, provided powerful identifying restrictions. In the several decades since the dawn of the structural micro-econometric approach, some insights from behavioral economics and lab experiments have been gaining more acceptance within mainstream economics, and it would seem as if it would be to the structural approach's advantage to judiciously adopt them. It is necessary to proceed with caution in this transition, though, as behavioral models impose fewer restrictions on possible observed behavior and therefore, create additional identification challenges. See Rust (2019) for ideas along these lines and DellaVigna (2018) for a survey of this new "structural behavioral economics" approach.

- **Machine Learning.** We also believe that there is much scope for structural models in many applied microeconomic fields to integrate with newly popular machine learning techniques in the same way consumer choice models in marketing and industrial organization and causal inference approaches are already integrating. Iskhakov et al. (2020) provide a survey of this new literature that explores possible synergies between structural econometrics and machine learning.

- **Bounds.** In some settings it might be useful for structural approaches absorb insights from the literature on bounds and partial identification as in the work of Manski (2014) and Kline and Tartari (2016). While this approach has the advantage of being rather agnostic and aims to impose only the minimal assumptions necessary, it might often lead to inconclusive findings.[41]

- **Non-Parametrics.** while most structural work uses tightly parameterized models, much credibility can be gained by being more non-parametric in functional form or distributional assumptions. Non-parametric analysis is worth doing at least to show that identification of the structure does not rely on those assumptions, even though in the end, due to data limitations, actual estimation might proceed using parametric restrictions. If feasible, non-parametric estimation might be particularly useful when the counterfactuals or environmental changes of interest are within the support of the data. Matzkin (1994, 2007, 2013) provide general treatments on non-parametric identification of structural models.

[41] For a pessimistic view on this partial identification approach to structural models or what he terms "nothing in, nothing out", see Rust (2016).

In conclusion, we view structural models as enjoying a resurgence in recent years with many avenues for gaining popularity moving forward. By tightly connecting theory and measurement, the structural approach remains a powerful empirical tool with the unique and sometimes exclusive capability to answer certain scientific and policy questions in microeconomics. We have provided what we hope to be a balanced introduction and overview to the structural approach, that highlights its advantages but also acknowledges some of its weaker fronts. We believe that many of the concerns some have with the approach can be ameliorated or eliminated by making the approach more receptive to new trends in empirical work in microeconomics, which emphasize the transparency of the research design and the sources of exogenous variation. We have then provided a road-map to understand how this powerful integration of model-based and design-based approaches has already begun to take place.

References

Aakvik, A., J. J. Heckman, and E. J. Vytlacil (1999). Semiparametric program evaluation: Lessons from an evaluation of a Norwegian training program. Unpublished manuscript, University of Chicago, Department of Economics.

Aakvik, A., J. J. Heckman, and E. J. Vytlacil (2005). Estimating treatment effects for discrete outcomes when responses to treatment vary: An application to Norwegian vocational rehabilitation programs. *Journal of Econometrics 125*(1–2), 15–51.

Abbring, J. (2010). Identification of dynamic discrete choice models. *Annual Reviews of Economics 2*, 367–394.

Abbring, J. (2012). Mixed hitting-time models. *Econometrica 80*(2), 783–819.

Abbring, J. and N. Daljord (2020). Identifying the discount factor in dynamic discrete choice models. *Quantitative Economics 11*(2), 471—-501.

Ackerberg, D. (2009). A new use of importance sampling to reduce computational burden in simulation estimation. *Quantitative Marketing and Economics 7*, 343–376.

Ackerberg, D., L. Benkard, S. Berry, and A. Pakes (2007). Econometric tools for analyzing market outcomes. In J. J. Heckman and E. E. Leamer (Eds.), *Handbook of Econometrics, Volume 6A*, Chapter 63, pp. 4171–4276. Amsterdam: Elsevier B. V.

Adda, J. and R. Cooper (2003). *Dynamic Economics*. Princeton: The MIT Press.

Adda, J. and C. Dustmann (2021). Sources of wage growth. *Journal of Political Economy (Forthcoming)*.

Adda, J., C. Dustmann, and K. Stevens (2017). The career costs of children. *Journal of Political Economy 125*(2), 293–337.

Agostinelli, F., E. Borghesan, and G. Sorrenti (2020). Welfare,workfare and labor supply: A unified ex post and ex ante evaluation. Working Paper.

Agostinelli, F. and M. Wiswall (2020). Estimating the technology of children's skill formation. NBER Working Paper No. 22442.

Aguirregabiria, V. and P. Mira (2002). Swapping the nested fixed point algorithm: A class of estimators for discrete markov decision models. *Econometrica 70*, 1519–1543.

Aguirregabiria, V. and P. Mira (2010). Dynamic discrete choice structural models: A survey. *Journal of Econometrics 156*(1), 38–67.

Altug, S. and R. A. Miller (1998). The effect of work experience on female wages and labour supply. *Review of Economic Studies 65*(1), 45–85.

Andrews, I., M. Gentzkow, and J. M. Shapiro (2017). Measuring the sensitivity of parameter estimates to estimation moments. *Quarterly Journal of Economics 132*(4), 1553–1592.

Andrews, I., M. Gentzkow, and J. M. Shapiro (2020a). On the informativeness of descriptive statistics for structural estimates. *Econometrica 88*(6), 2231—-2258.

Andrews, I., M. Gentzkow, and J. M. Shapiro (2020b). Transparency in structural research. *Journal of Business and Economic Statistics 38*(4), 711–722.

Angrist, J. and J. Pischke (2010). The credibility revolution in empirical economics: How better research design is taking the con out of econometrics. *Journal of Economic Perspectives 24*(2), 3–30.

Arcidiacono, P. (2004). Ability sorting and the returns to college major. *Journal of Econometrics 121*, 343–375.

Arcidiacono, P., P. Bayer, J. Blevins, and P. Ellickson (2016). Estimation of dynamic discrete choice models in continuous time with an application to retail competition. *Review of Economic Studies 83*, 889–931.

Arcidiacono, P. and P. B. Ellickson (2011). Practical methods for estimation of dynamic discrete choice models. *Annual Review of Economics 3*(1), 363–394.

Arcidiacono, P. and J. B. Jones (2003). Finite mixture distributions, sequential likelihood and the em algorithm. *Econometrica 71*(3), 933–946.

Arcidiacono, P. and R. Miller (2020). Identifying dynamic discrete choice models off short panels. *Journal of Econometrics 215*(2), 473–485.

Arcidiacono, P. and R. A. Miller (2011). Conditional choice probability estimation of dynamic discrete choice models with unobserved heterogeneity. *Econometrica 7*(6), 1823–1868.

Arcidiacono, P., K. Muralidharan, E.-y. Shim, and J. D. Singleton (2021). Experimentally validating welfare evaluation of school vouchers: Part i. National Bureau of Economic Research, Working Paper No. 29077.

Arcidiacono, P., H. Sieg, and F. Sloan (2007). Living rationally under the volcano? an empirical analysis of heavy drinking and smoking. *International Economic Review 48*(1), 37–65.

Athey, S. and P. Haile (2007). Nonparametric approaches to auctions. In J. J. Heckman and E. E. Leamer (Eds.), *Handbook of Econometrics Vol 6A*, Chapter 60, pp. 3847–3965. Amsterdam: Elsevier B. V.

Attanasio, O., C. Meghir, and A. Santiago (2012). Education choices in mexico: Using a structural model and a randomized experiment to evaluate progresa. *Review of Economic Studies 79*(1), 37–66.

Avery, R., L. P. Hansen, and V. J. Hotz (1983). Multiperiod probit models and orthogonality condition estimation. *International Economic Review 24*(1), 21–35.

Bajari, P., L. Benkard, and J. Levin (2007). Estimating dynamic models of imperfect competition. *Econometrica 75*(5), 1331–1370.

Bajari, P., C. Sean Chu, D. Nekipelov, and M. Park (2016). Identification and semiparametric estimation of a finite horizon dynamic discrete choice model with a terminating action. *Quantitative Marketing and Economics 14*(4), 271–323.

Banerjee, A. and E. Duflo (2009). The experimental approach to development economics. *Annual Review of Economics 1*, 151—-78.

Belzil, C. (2007). The return to schooling in structural dynamic models: a survey. *European Economic Review 51*, 1059–1105.

Bergman, P., R. Chetty, S. DeLuca, N. Hendren, L. F. Katz, and C. Palmer (2020). Creating moves to opportunity: Experimental evidence on barriers to neighborhood choice. Working Paper.

Berkovec, J. and S. Stern (1991). Job exit behavior of older men. *Econometrica 59*(1), 189–210.

Blau, D. and D. Gilleskie (2001). Retiree health insurance and the labor force behavior of older men in the 1990s. *The Review of Economics and Statistics 83*(1), 64–80.

Blevins, J. R. (2014). Non-parametric identification of dynamic decision processes with discrete and continuous choices. *Quantitative Economics 5*, 531–554.

Blundell, R. (2017). What have we learned from structural models? *American Economic Review 107*(5), 287–292.

Blundell, R., M. Costa Dias, C. Meghir, and J. Shaw (2016). Female labor supply, human capital and welfare reform. *Econometrica 84*(5), 1705–1753.

Blundell, R., A. Duncan, and C. Meghir (1998). Estimating labor supply responses using tax reforms. *Econometrica 66*(4), 827–861.

Blundell, R. and T. Macurdy (1999). Labor supply: A review of alternative approaches. In O. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics, Volume 3*, Chapter 27, pp. 1559–1695. Amsterdam: Elsevier.

Blundell, R., T. Macurdy, and C. Meghir (2007). Labor supply models: Unobserved heterogeneity, non-participation and dynamics. In J. J. Heckman and E. Leamer (Eds.), *Handbook of Econometrics, Volume 6A*, Chapter 69, pp. 4669–4775. Amsterdam: Elsevier B. V.

Bo, H. and S. Galiani (2020). Assesing external validity. NBER Working Paper No. 26422.

Bobba, M., L. Flabbi, and S. Levy (2021). Labor market search, informality and schooling investments. *International Economic Review (Forthcoming).*

Bonhomme, S., T. Lamadon, and E. Manresa (2021). Discretizing unobserved heterogeneity. Working Paper.

Brien, M. J., L. A. Lillard, and S. Stern (2006). Cohabitation, marriage and divorce in a model of match quality. *International Economic Review 47*(2), 451–494.

Bruins, M., J. A. Duffy, M. P. Keane, and A. A. Smith (2018). Generalized indirect inference for discrete choice models. *Journal of Econometrics 205*(1), 177–203.

Burtless, G. and J. Hausman (1978). The effect of taxation on labor supply: Evaluating the gary negative income tax experiment. *Journal of Political Economy 86*(6), 1103–1130.

Cameron, S. V. and J. J. Heckman (1987). Son of ctm: The dcpa approach based on discrete and factor structure models. Unpublished working paper, University of Chicago, Department of Economics.

Canals, J. J. and S. Stern (2002). Empirical search models. In S. Woodbury and C. Davidson (Eds.), *Search Theory and Unemployment*. Kluwer Academic Publishers.

Card, D. and D. Hyslop (2005). Estimating the effects of a time-limited earnings subsidy for welfare leavers. *Econometrica 6*(73), 1723–1770.

Carneiro, P., K. Hansen, and J. J. Heckman (2003). Estimating distributions of treatment effects with an application to the returns to schooling and measurement of the effects of uncertainty on college choice. *International Economic Review 44*(2), 361–422.

Chan, M. (2017). Welfare dependence and self-control: An empirical analysis. *Review of Economic Studies 84*, 1379–1423.

Chan, M. and R. Moffitt (2018). Welfare reform and the labor market. *Annual Review of Economics 10*, 347–381.

Chaparro, J., A. Sojourner, and M. Wiswall (2020). Early childhood care and cognitive development. NBER Working Paper No. 26813.

Chetty, R. (2009). Sufficient statistics for welfare analysis: A bridge between structural and reduced-form methods. *Annual Review of Economics 1*, 451–488.

Chiappori, P.-A. (1988). Rational household labor supply. *Econometrica 56*(1), 63–90.

Chiappori, P.-A., M. Costa-Dias, and C. Meghir (2019). The marriage market, labor supply and education choice. *Journal of Political Economy 56*(1), 63–90.

Chiappori, P.-A. and M. Mazzoco (2017). Static and intertemporal household decisions. *Journal of Economic Literature 55*(3), 985–1045.

Choi, E. J. (2018). Evaluating a structural model of labor supply and welfare participation: Evidence from state welfare reform experiments. Working Paper, Hanyang University.

Choo, E. and A. Siow (2006). Who marries whom and why. *Journal of Political Economy 114*(1), 175–201.

Christensen, B. J. and N. M. Kiefer (2009). *Economic Modeling and Inference*. Princeton: Princeton University Press.

Cornelissen, T., C. Dustmann, A. Raute, and U. Schönberg (2016). From late to mte: Alternative methods for the evaluation of policy interventions. *Labour Economics 41*, 47–60.

Cunha, F. and J. J. Heckman (2008). Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation. *Journal of Human Resources 43*(4), 738–782.

Cunha, F., J. J. Heckman, and S. M. Schennach (2010). Estimating the technology of cognitive and noncognitive skill formation. *Econometrica 78*(3), 883–931.

Del Boca, D. and C. Flinn (2012). Endogenous household interaction. *Journal of Econometrics 166*(1), 49–65.

Del Boca, D., C. Flinn, and M. Wiswall (2014). Household choices and child development. *Review of Economic Studies 81*(137–185), 49–65.

DellaVigna, S. (2018). Structural behavioral economics. In D. Bernheim, S. DellaVigna, and D. Laibson (Eds.), *Handbook of Behavioral Economics, Foundations and Applications 1*, Chapter 7, pp. 613–723. Amsterdam: Elsevier B.V.

Devine, T. and N. Kiefer (1991). *Empirical Labor Economics. The Search Approach*. New York: Oxford University Press.

Dey, M. S. and C. Flinn (2005). An equilibrium model of health insurance provision and wage determination. *Econometrica 73*(2), 571–627.

Dey, M. S. and C. Flinn (2008). Household search and health insurance coverage. *Journal of Econometrics 145*, 43–63.

d'Haultfoeuille, X. and P. Février (2020). The provision of wage incentives: A structural estimation using contracts variation. *Quantitative Economics 11*, 349–397.

Duflo, E., R. Hanna, and S. Ryan (2012). Incentives work: Getting teachers to come to school. *American Economic Review 102*(4), 1241–1278.

Eberwein, C. and J. Ham (2008). Obtaining analytic derivatives for a popular discrete-choice dynamic programming model. *Economic Letters 101*, 168–171.

Eckstein, Z., M. Keane, and O. Lifshitz (2019). Career and family decisions: Cohorts born 1935-1975. *Econometrica 87*(1), 217—-253.

Eckstein, Z. and G. J. van den Berg (2007). Empirical labor search: A survey. *Journal of Econometrics 136*, 531–564.

Eckstein, Z. and K. I. Wolpin (1989a). Dynamic labour force participation of married women and endogenous work experience. *The Review of Economic Studies 56*, 375–390.

Eckstein, Z. and K. I. Wolpin (1989b). The specification and estimation of dynamic stochastic discrete choice models: A survey. *The Journal of Human Resources 24*(4), 562–598.

Eckstein, Z. and K. I. Wolpin (1990). On the estimation of labour force participation, job search and job matching models using panel data. In Weiss and Fischelson (Eds.), *Advances in the Theory and Measurement of Unemployment*, pp. 82–112. MacMillan.

Eckstein, Z. and K. I. Wolpin (1999). Why youths drop out of high school: The impact of preferences, opportunities, and abilities. *Econometrica 67*(6), 1295–1339.

Einav, L. and A. Finkelstein (2018). Moral hazard in health insurance: What do we know and how do we know it? *Journal of the European Economic Association 16*(4), 957–982.

Einav, L. and J. Levin (2010). Empirical industrial organization: A progress report. *Journal of Economic Perpectives 24*(2), 145–162.

Eisenhauer, P., J. Gabler, and L. Janys (2021). Structural models for policy-making. coping with parametric uncertainty. Working Paper, University of Bonn.

Eisenhauer, P., J. J. Heckman, and S. Mosso (2015). Estimation of dynamic discrete choice models by maximum likelihood and the simulated method of moments. *International Economic Review 56*(2: Special Issue: Estimation of Dynamic Stochastic Models in Empirical Microeconomics: In Honor of Kenneth I. Wolpin), 331–357.

Falk, A., A. Becker, T. Dohmen, B. Enke, D. Huffman, and U. Sunde (2018). Global evidence on economic preferences. *The Quarterly Journal of Economics 133*(4), 1645–1692.

Fang, H. and D. Silverman (2009). Time-inconsistency and welfare program participation: Evidence from the nlsy. *International Economic Review 50*(4), 1043–1077.

Fang, H. and Y. Wang (2015). Estimating dynamic discrete choice models with hyperbolic discounting, with an application to mammography decisions. *International Economic Review 56*(2), 565–596.

Ferrall, C. (2012). Explaining and forecasting results of the self-sufficiency project. *Review of Economic Studies 79*, 1495–1526.

Ferrall, C. and B. Shearer (1999). Incentives and transaction costs within the firm: Estimating an agency model using payroll records. *Review of Economic Studies 99*, 309–338.

Ferreyra, M. M. (2007). Estimating the effects of private school vouchers in multidistrict economies. *American Economic Review 97*(3), 789–817.

Flinn, C. (2010). *The Minimum Wage and Labor Market Outcomes*. Cambridge: The MIT Press.

Flinn, C. and J. Heckman (1982). New methods for analyzing structural models of labor force dynamics. *Journal of Econometrics 18*, 115–168.

Fox, J., K.-i. Kim, S. Ryan, and P. Bajari (2011). A simple estimator for the distribution of random coefficients. *Quantitative Economics 2*, 381–418.

Frederick, S., G. Loewenstein, and T. O'Donoghue (2002). Time discounting and time preference: A critical review. *Journal of Economic Literature XL*, 351–401.

French, E. (2005). The effect of health, wealth, and wages on labour supply and retirement behaviour. *Review of Economic Studies 72*, 395–427.

French, E. and J. B. Jones (2011). The effects of health insurance and self-insurance on retirement behavior. *Econometrica 79*(3), 693–732.

French, E. and J. B. Jones (2017). Health, health insurance, and retirement: A survey. *Annual Reviews of Economics 9*, 383–409.

French, E. and C. Taber (2011). Identification of models of the labor market. In O. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics*, Volume 4A, Chapter 6, pp. 537–617. Elsevier Science B. V.

Galiani, S., A. Murphy, and J. Pantano (2015). Estimating neighborhood choice models: Lessons from a housing assistance experiment. *American Economic Review 105*(11), 3385–3415.

Gayle, G.-L., L. Golan, and R. Miller (2015). Promotion, turnover and compensation in the executive labor market. *Econometrica 83*(6), 2293–2369.

Gayle, G.-L., L. Golan, and M. Soytas (2018). Estimation of dynastic life-cycle discrete choice models. *Quantitative Economics 9*, 1195–1241.

Gayle, W.-R. (2018). Ccp estimation of dynamic discrete/continuous-choice models with generalized finite dependence and correlated unobserved heterogeneity. Working Paper.

Gentry, M. L., T. P. Hubbard, D. Nekipelov, and H. J. Paarsch (2018). Structural econometrics of auctions: A review. *Foundations and Trends in Econometrics 9*(2-4), 79–302.

Geweke, J. and M. Keane (2001). Computationally intensive methods of integration in econometrics. In J. J. Heckman and E. E. Leamer (Eds.), *Handbook of Econometrics*, Volume 5, Chapter 56, pp. 3463–3568. Amsterdam: Elsevier Science B. V.

Gotz, G. A. and J. J. McCall (1984). A dynamic retention model for air force officers. theory and estimates. Technical report, Rand Corporation report.

Gourieroux, C., A. Monfort, and E. Renault (1993). Indirect inference. *Journal of Applied Econometrics 8*((S1)), S85–S118.

Gousse, M., N. Jacquemet, and J.-M. Robin (2019). Marriage, labor supply and home production. *Econometrica 85*(6), 1873–1919.

Grogger, J. and L. Karoly (2005). *Welfare Reform. Effects of a Decade of Change*. Cambridge, MA: Harvard University Press.

Ham, J. and T. Woutersen (2011). Calculating confidence intervals for continuous and discontinuous functions of estimated parameters. IZA Working Paper.

Heckman, J., L. Lochner, and C. Taber (1998). Explaining rising wage inequality: Explorations with a dynamic equilibrium model of labor earnings with heterogeneous agents. *Review of Economic Dynamics 1*, 1–58.

Heckman, J., L. Lochner, and P. Todd (2006). Earnings functions, rates of return and treatment effects: The mincer equation and beyond. In E. A. Hanushek and F. Welch (Eds.), *Handbook of the Economics of Education, Volume 1*, Chapter 7, pp. 307–458. Amsterdam: Elsevier B. V.

Heckman, J. and S. Urzua (2010). Comparing iv with structural models: What simple iv can and cannot identify. *Journal of Econometrics 156*, 27–37.

Heckman, J. J. (1974). Shadow prices, market wages, and labor supply. *Econometrica 42*(4), 679–694.

Heckman, J. J. (1981a). The incidental parameters problem and the problem of initial conditions in estimating a discrete time-discrete data stochastic process and some Monte Carlo evidence. In C.Manski and D. McFadden (Eds.), *Structural Analysis of Discrete Data With Econometric Applications*, pp. 179–195. Cambridge, MA: MIT Press.

Heckman, J. J. (1981b). Statistical models for discrete panel data. In C. Manski and D. McFadden (Eds.), *Structural Analysis of Discrete Data with Econometric Applications*, pp. 114–178. Cambridge, MA: MIT Press.

Heckman, J. J. (1992). Randomization and social policy evaluation. In C. F. Manski and I. Garfinkel (Eds.), *Evaluating Welfare and Training Programs*, Chapter 5, pp. 201–230. Cambridge, MA: Harvard University Press.

Heckman, J. J. (1997). A study of implicit behavioral assumptions used in making program evaluations. *The Journal of Human Resources 32*(3), 441–462.

Heckman, J. J. (2010). Building bridges between structural and program evaluation approaches to evaluating policy. *Journal of Economic Literature 48*, 356–398.

Heckman, J. J. (2020). Randomization and social policy evaluation revisited. HCEO Working Paper Series.

Heckman, J. J. and B. S. Singer (1984). A method for minimizing the impact of distributional assumptions in econometric models for duration data. *Econometrica 52*(2), 271–320.

Heckman, J. J. and J. Smith (1995). Assessing the case for social experiments. *Journal of Economic Perspectives 9*(2), 85–110.

Heckman, J. J. and E. Vytlacil (2005). Structural equations, treatment effects, and econometric policy evaluation. *Econometrica 73*(3), 669–738.

Heckman, J. J. and E. J. Vytlacil (1999). Local instrumental variables and latent variable models for identifying and bounding treatment effects. *Proceedings of the National Academy of Sciences 96*(8), 4730–4734.

Heckman, J. J. and E. J. Vytlacil (2001). Policy-relevant treatment effects. *American Economic Review 91*(2), 107–111.

Holland, P. W. (1986). Statistics and causal inference. *Journal of the American Statistical Association 81*(396), 945–960.

Honoré, B. E., T. H. Jørgensen, and A. de Paula (2020). The informativeness of estimation moments. *Journal of Applied Econometrics 35*(7), 797–813.

Hotz, V., J. Klerman, and R. Willis (1997). The economics of fertility in developed countries. In M. Rosenzweig and O. Stark (Eds.), *Handbook of Population and Family Economic*, Volume 1, Chapter 7, pp. 275–347. Amsterdam: Elsevier Science B. V.

Hotz, V. J., R. Miller, S. Sanders, and J. Smith (1994). A simulation estimator for dynamic models of discrete choice. *Review of Economic Studies 61*, 265–289.

Hotz, V. J. and R. A. Miller (1993). Conditional choice probabilities and the estimation of dynamic models. *Review of Economic Studies 60*(3), 497–529.

Hu, Y. and M. Shum (2012). Nonparametric identification of dynamic models with unobserved state variables. *Journal of Econometrics 171*(1), 32–44.

Hurwicz, L. (1962). On the structural form of interdependent systems. In E. Nagel, P. Suppes, and A. Tarski (Eds.), *Logic, Methodology and Philosophy of Science*, pp. 232–239. Stanford: Stanford University Press.

Ichimura, H. and C. Taber (2002). Semiparametric reduced-form estimation of tuition subsidies. *American Economic Review 92*(2), 286–292.

Imai, S., N. Jain, and A. Ching (2009). Bayesian estimation of dynamic discrete choice models. *Econometrica 77*(6), 1865—-1899.

Imai, S. and M. Keane (2004). Intertemporal labor supply and human capital accumulation. *International Economic Review 45*(2), 601–641.

Imbens, G. (2010). Better late than nothing: Some comments on deaton (2009) and heckman and urzua (2009). *Journal of Economic Literature 48*, 399–423.

Imbens, G., D. Rubin, and B. Sacerdote (2001). Estimating the effect of unearned income on labor earnings, savings, and consumption: Evidence from a survey of lottery players. *American Economic Review 91*(4), 778–794.

Imbens, G. W. and J. D. Angrist (1994). Identification and estimation of local average treatment effects. *Econometrica 62*(2), 467–475.

Iskhakov, F., T. H. Jørgensen, J. Rust, and B. Schjerning (2017). The endogenous grid method for discrete-continuous dynamic choice models with (or without) taste shocks. *Quantitative Economics 8*(2), 317–365.

Iskhakov, F., J. Rust, and B. Schjerning (2020). Machine learning and structural econometrics: Contrasts and synergies. *Economic Journal 23*(3), S81–S124.

Jørgensen, T. H. (2021). Sensitivity to calibrated parameters. *Review of Economics and Statistics (forthcoming)*.

Johnson, M. and M. Keane (2013). A dynamic equilibrium model of the us wage structure, 1968–1996. *Journal of Labor Economics 31*(1), 1–49.

Kaji, T., E. Manresa, and G. Pouliot (2020). An adversarial approach to structural estimation. Working Paper.

Kalouptsidi, M., P. T. Scott, and E. Souza-Rodrigues (2021). Identification of counterfactuals in dynamic discrete choice models. *Quantitative Economics 12*(2), 351–403.

Kasahara, H. and K. Shimotsu (2009). Non-parametric identification of finite mixture models of dynamic discrete choices. *Econometrica 77*(1), 135–175.

Keane, M. (1992). A note on identification in the multinomial probit model. *Journal of Business & Economic Statistics 10*(2), 193–200.

Keane, M. (2010a). A structural perspective on the experimentalist school. *Journal of Economic Perspectives 24*(2), 47–58.

Keane, M. (2010b). Structural vs. atheoretic approaches to econometrics. *Journal of Econometrics 156*, 3–20.

Keane, M. (2011). Labor supply and taxes: A survey. *Journal of Economic Literature 49*(4), 961–1075.

Keane, M. (2015). Effects of permanent and transitory tax changes in a life-cycle labor supply model with human capital. *International Economic Review 56*(2), 485–503.

Keane, M. and R. Moffitt (1998). A structural model of multiple welfare participation and labor supply. *International Economic Review 39*(3), 553–589.

Keane, M., P. Todd, and K. I. Wolpin (2011). The structural estimation of behavioral models: Discrete choice dynamic programming methods and applications. In O. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics*, Volume 4A, Chapter 4, pp. 331–461. Amsterdam: Elsevier Science B. V.

Keane, M. and K. I. Wolpin (1994). The solution and estimation of discrete choice dynamic programming models by simulation and interpolation: Montecarlo evidence. *Review of Economics and Statistics 76*(4), 648–672.

Keane, M. and K. I. Wolpin (1997). The career decisions of young men. *Journal of Political Economy 105*(3), 473–522.

Keane, M. and K. I. Wolpin (2001). The effect of parental transfers and borrowing constraints on educational attainment. *International Economic Review 42*(4), 1051–1103.

Keane, M. and K. I. Wolpin (2007). Exploring the usefulness of a nonrandom holdout sample for model validation: Welfare effects on female behavior. *International Economic Review 48*(2), 1351–1378.

Keane, M. and K. I. Wolpin (2009). Empirical applications of discrete choice dynamic programming models. *Review of Economic Dynamics 12*, 1–22.

Kline, P. and M. Tartari (2016). Bounding the labor supply responses to a randomized welfare experiment: A revealed preference approach. *American Economic Review 4*(106), 972–1014.

Kline, P. and C. Walters (2019). On heckit, late and numerical equivalence. *Econometrica 87*(2), 677–696.

Kristensen, D., L. Nesheim, and A. de Paula (2015). Ccp and the estimation of nonseparable dynamic models. Working Paper.

Lalonde, R. (1986). Evaluating the econometric evaluations of training programs with experimental data. *American Economic Review 76*(4), 604–620.

Lee, D. (2005). An estimable dynamic general equilibrium model of work, schooling and occupational choice. *International Economic Review 46*(1), 1–34.

Lee, D. and K. I. Wolpin (2006). Intersectoral labor mobility and the growth of the service sector. *Econometrica 74*(1), 1–46.

Lee, D. and K. I. Wolpin (2009). Accounting for wage and employment changes in the us from 1968-2000: A dynamic model of labor market equilibrium. *Journal of Econometrics 156*(1), 68–85.

Lerman, S. and C. Manski (1981). On the use of simulated frequencies to approximate choice probabilities. In C. Manski and D. McFadden (Eds.), *Structural Analysis of Discrete Data and Econometric Applications*, pp. 305–319. Cambridge, MA: MIT Press.

Levy, M. R. and P. Schiraldi (2020). Identification of intertemporal preferences in history-dependent dynamic discrete choice models. Working Paper.

Lise, J., S. Seitz, and J. Smith (2015). Evaluating search and matching models using experimental data. *IZA Journal of Labor Economics 4*(16), 2–35.

List, J. (2020). Non est disputandum de generalizability? a glimpse into the external validty trial. NBER Working Paper No. 27535.

List, J. and I. Rasul (2011). Field experiments in labor economics. In O. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics, Volume 4a*, Chapter 2, pp. 103–228. Amsterdam: Elsevier B. V.

Low, H. and C. Meghir (2017). The use of structural models in econometrics. *Journal of Economic Perspectives 31*(2), 33–58.

MaCurdy, T. (1981). An empirical model of labor supply in a life cycle setting. *Journal of Political Economy 6*(89), 1059–1085.

Magnac, T. and D. Thesmar (2002). Identifying dynamic discrete decision processes. *Econometrica 70*(2), 801–816.

Mahajan, A., C. Michel, and A. Tarozzi (2020). Identification of time-inconsistent models: The case of insecticide treated nets. Working Paper.

Manski, C. (2008). Measuring expectations. *Econometrica 72*(5), 1329–1376.

Manski, C. (2014). Identification of income-leisure preferences and evaluation of income tax policy. *Quantitative Economics 5*, 145–174.

Margiotta, M. and R. Miller (2000). Managerial compensation and the costs of moral hazard. *International Economic Review 41*(3), 669–719.

Marschak, J. (1953). Economic measurements for policy and prediction. In W. C. Hood and T. C. Koopmans (Eds.), *Studies in Econometric Method*, Chapter 1, pp. 1–26. New York: Wiley.

Matzkin, R. (1994). Restrictions of economic theory in non-parametric methods. In R. F. Engle and D. L. McFadden (Eds.), *Handbook of Econometrics, Volume 4*, Volume 4, Chapter 42, pp. 2523–2558. New York, NY: Elsevier B. V.

Matzkin, R. (2007). Nonparametric identification. In E. E. L. James J. Heckman (Ed.), *Handbook of Econometrics Volume 6B*, Chapter 73, pp. 5307–5368. Amsterdam: Elsevier B. V.

Matzkin, R. (2013). Non-parametric identification in structural economic models. *Annual Review of Economics 5*, 457–486.

McFadden, D. (1977). Validation of disaggregate travel demand models: Some tests. urban demand forecasting project. final report. volume v. Technical report, Institute of Transportation Studies, University of California, Berkeley.

McFadden, D. (1989). A method of simulated moments for estimation of discrete response models without numerical integration. *Econometrica 84*(57), 995–1026.

Meghir, C. (2006). Dynamic models for policy evaluation. In R. Blundell, W. K. Newey, and T. Persson (Eds.), *Advances in Economics and Econometrics, Theory and Applications, Ninth World Congress*. Cambridge University Press.

Miller, R. A. (1984). Job matching and occupational choice. *Journal of Political Economy 92*(6), 1086–1120.

Miller, R. A. (1999). *Estimating Models of Dynamic Optimization with Microeconomic Data*, Chapter 6, pp. 227–274. John Wiley and Sons, Ltd.

Mira, P. (2007). Uncertain infant mortality, learning and life-cycle fertility. *International Economic Review 48*(3), 809–846.

Misra, S. and H. Nair (2011). A structural model of sales-force compensation dynamics: Estimation and field implementation. *Quantitative Marketing and Economics* -(9), 211–257.

Moffitt, R. (1983). An economic model of welfare stigma. *American Economic Review 73*(5), 1023–1035.

Moffitt, R. A. (1979). The labor supply response in the gary experiment. *Journal of Human Resources 14*(4), 477–487.

Mogstad, M., A. Torgovitsky, and C. Walters (2020). The causal interpretation of two-stage least squares with multiple instrumental variables. NBER Working Paper 25691.

Mroz, T. A. (1999). Discrete factor approximations in simultaneous equation models: Estimating the impact of a dummy endogenous variable on a continuous outcome. *Journal of Econometrics 92*, 233–274.

Navarro, S. (2022). Selection models and control functions. In *Handbook of Labor, Population and Human Resources*. Springer.

Nevo, A. and M. Whinston (2010). Taking the dogma out of econometrics: Structural modeling and credible inference. *Journal of Economic Perpectives 24*(2), 69–82.

Norets, A. (2009). Inference in dynamic discrete choice models with serially correlated unobserved state variables. *Econometrica 77*(5), 1665–1682.

Norets, A. and X. Tang (2014). Semiparametric inference in dynamic binary choice models. *The Review of Economic Studies 81*(3), 1229–1262.

Paarsch, H. and B. Shearer (2000). Piece rates, fixed wages and incentive effects: Statistical evidence from payroll records. *International Economic Review 41*, 59–92.

Pakes, A. (1986). Patents as options: Some estimates of the value of holding european patent stocks. *Econometrica 54*(4), 755–784.

Pantano, J. and Y. Zheng (2013). Using subjective expectation data to allow for unobserved heterogeneity in hotz-miller estimation strategies. Working Paper.

Pistaferri, L. (2003). Anticipated and unanticipated wage changes, wage risk, and intertemporal labor supply. *Journal of Labor Economics 21*(3), 729–754.

Postel Vinay, F. and J.-M. Robin (2006). Microeconometric search-matching models and matched employer-employee data. In R. Blundell, W. K. Newey, and T. Persson (Eds.), *Advances in Economics and Econometrics, Theory and Applications, Ninth World Congress*. Cambridge University Press.

Prendergast, C. (1999). The provision of incentives in firms. *Journal of Economic Literature 37*, 7–63.

Reiss, P. C. and F. A. Wolak (2007). Structural econometric modeling: Rationales and examples from industrial organization. In J. J. Heckman and E. E. Leamer (Eds.), *Handbook of Econometrics, Volume 6A*, Chapter 64, pp. 4277–4415. Amsterdam: Elsevier B.v.

Rosenzweig, M. and K. I. Wolpin (2000). Natural "natural experiments" in economics. *Journal of Economic Literature 38*, 827–874.

Rust, J. (1987). Optimal replacement of GMC bus engines: An empirical model of Harold Zurcher. *Econometrica 55*(5), 999–1033.

Rust, J. (1994). Structural estimation of Markov decision processes. In R. F. Engle and D. L. McFadden (Eds.), *Handbook of Econometrics*, Volume 4, Chapter 51, pp. 3081–3143. New York, NY: North-Holland.

Rust, J. (1996). Numerical dynamic programming in economics. In H. Amman, D. Kendrick, and J. Rust (Eds.), *Handbook of Computational Economics*, Volume 1, Chapter 14, pp. 619–729. Amsterdam: Elsevier Science B. V.

Rust, J. (2010). Comments on: "structural vs. atheoretic approaches to econometrics" by michael keane. *Journal of Econometrics 156*, 21–24.

Rust, J. (2014). The limits of inference with theory: A review of wolpin (2013). *Journal of Economic Literature 3*(52), 820–850.

Rust, J. (2016). Mostly useless econometrics? assessing the causal effect of econometric theory. *Foundations and Trends in Accounting 10*(2-4), 127–203.

Rust, J. (2019). Has dynamic programming improved decision making? *Annual Review of Economics 11*, 833–858.

Sauer, R. (2004). Educational financing and lifetime earnings. *The Review of Economic Studies 71*(4), 1189–1216.

Sauer, R. M. and C. R. Taber (2021). Understanding women's wage growth using indirect inference with importance sampling. *Journal of Applied Econometrics 36*(4), 453–473.

Schorfheide, F. and K. I. Wolpin (2012). On the use of holdout samples for model selection. *American Economic Review: Papers and Proceedings 102*(3), 477–481.

Schorfheide, F. and K. I. Wolpin (2016). To hold out or not to hold out. *Research in Economics 70*, 332–345.

Shephard, A. (2019). Marriage market dynamics, gender, and the age gap. Working Paper, University of Pennsylvanya.

Stern, S. (1989). Estimating a simultaneous search model. *Journal of Labor Economics 7*(3), 348–369.

Stern, S. (1992). A method for smoothing simulated moments of discrete probabilities in multinomial probit models. *Econometrica 60*(4), 943–952.

Stern, S. (1997). Simulation-based estimation. *Journal of Economic Literature 35*(4), 2006–2039.

Su, C. L. and K. Judd (2012). Constrained optimization approaches to estimation of structural models. *Econometrica 80*(5), 2213–2230.

Taber, C. (2000). Semiparametric identification and heterogeneity in discrete choice dynamic programming models. *Journal of Econometrics 96*, 201–229.

Taber, C. R. and R. Vejlin (2020). Estimation of a roy/search/compensating differential model of the labor market. *Econometrica 88*(3), 1031–1069.

Tartari, M. (2015). Divorce and the cognitive ability of children. *International Economic Review 56*(2), 597–645.

Todd, P. and K. I. Wolpin (2006). Assessing the impact of a school subsidy program in mexico: Using a social experiment to validate a dynamic behavioral model of child schooling and fertility. *American Economic Review 5*(96), 1384–1417.

Todd, P. and K. I. Wolpin (2010). Structural estimation and policy evaluation in developing countries. *Annual Review of Economics 2*, 21–50.

Todd, P. and K. I. Wolpin (2020). The best of both worlds: Combining rcts with structural modeling. Working Paper.

van den Berg, G. (1990). Nonstationarity in job search theory. *The Review of Economic Studies 57*, 255–277.

van der Klaauw, W. (2012). On the use of expectations data in estimating structural dynamic choice models. *Journal of Labor Economics 30*(3), 521–554.

van der Klaauw, W. and K. I. Wolpin (2008). Social security and the retirement and savings behavior of low-income households. *Journal of Econometrics 145*, 21–42.

Voena, A. (2015). Yours, mine, and ours: Do divorce laws affect the intertemporal behavior of married couples? *American Economic Review 2015 8*(105), 2295–2332.

Willis, R. and S. Rosen (1979). Education and self-selection. *Journal of Political Economy 87*(5), S7–S36.

Wise, D. (1985). A behavioral model versus experimentation: the effects of housing subsidies on rent. In P. Brucker and R. Pauly (Eds.), *Methods of Operations Research*, Chapter 50, pp. 441–89. Konigstein: Verlag Anton Hain.

Wiswall, M. and B. Zafar (2015). Determinants of college major choice: Identification using an information experiment. *The Review of Economic Studies 82*(2), 791–824.

Wolpin, K. I. (1984). An estimable dynamic stochastic model of fertility and child mortality. *Journal of Political Economy 92*(5), 852–874.

Wolpin, K. I. (1987). Estimating a structural search model: The transition from school to work. *Econometrica 55*(4), 801–817.

Wolpin, K. I. (1995). *Empirical Methods for the Study of Labor Force Dynamics*. Luxembourg: Harwood Academic Publishers.

Wolpin, K. I. (1996). Public-policy uses of discrete-choice dynamic programming models. *American Economic Review 86*(2), 427–432.

Wolpin, K. I. (1997). Determinants and consequences of the mortality and health of infants and children. In M. Rosenzweig and O. Stark (Eds.), *Handbook of Population and Family Economic*, Volume 1, Chapter 7, pp. 483–557. Amsterdam: Elsevier Science B. V.

Wolpin, K. I. (2003). Wage equations and education policy. In M. Dewatripont, L. P. Hansen, and S. Turnovsky (Eds.), *Advances in Economics and Econometrics, Theory and Applications, Eight World Congress*. Cambridge University Press.

Wolpin, K. I. (2013). *The Limits of Inference Without Theory*. New York: The MIT Press.