

The Use of Structural Models in Econometrics

Hamish Low and Costas Meghir

The aim of this paper is to discuss the role of structural economic models in empirical analysis and policy design. This approach offers some valuable payoffs, but also imposes some costs.

Structural economic models focus on distinguishing clearly between the objective function of the economic agents and their opportunity sets as defined by the economic environment. The key features of such an approach at its best are a tight connection with a theoretical framework alongside a clear link with the data that will allow one to understand how the model is identified. The set of assumptions under which the model inferences are valid should be clear: indeed, the clarity of the assumptions is what gives value to structural models.

The central payoff of a structural econometric model is that it allows an empirical researcher to go beyond the conclusions of a more conventional empirical study that provides reduced-form causal relationships. Structural models define how outcomes relate to preferences and to relevant factors in the economic environment, identifying mechanisms that determine outcomes. Beyond this, they

■ *Hamish Low is Professor of Economics, Cambridge University, Cambridge, United Kingdom, and Cambridge INET and Research Fellow, Institute for Fiscal Studies, London, United Kingdom. Costas Meghir is the Douglas A. Warner III Professor of Economics, Yale University, New Haven, Connecticut; International Research Associate, Institute for Fiscal Studies, London, United Kingdom; Research Fellow, Institute of Labor Economics (IZA), Bonn, Germany; and Research Associate, National Bureau of Economic Research, Cambridge, Massachusetts. Their email addresses are hamish.low@econ.cam.ac.uk and c.meghir@yale.edu.*

† For supplementary materials such as appendices, datasets, and author disclosure statements, see the article page at <https://doi.org/10.1257/jep.31.2.33>

doi=10.1257/jep.31.2.33

are designed to analyze counterfactual policies, quantifying impacts on specific outcomes as well as effects in the short and longer run.

The short-run implications can often be compared to what actually happened in the data, allowing for validation of the model. For example, Blundell, Costa Dias, Meghir, and Shaw (2016) model how life-cycle female labor supply and human capital accumulation are affected by tax credit reform. They validate the model by comparing its short-run predictions to those estimated by simple reduced form methods. However, their model also has implications for labor supply and wages beyond the childbearing age, as well as for the educational choice of subsequent cohorts, none of which can be estimated from actual data without an economic model. Such effects are of central importance for understanding the impacts of welfare programs. Similarly, Low and Pistaferri (2015) model the long-run effects of reform to disability insurance, but validate their model using reduced form predictions. This symbiotic interaction of structural models and reduced form approaches, including randomized experiments, provides the strongest tool in the empirical economics toolkit and is emphasized in this paper.

Additional insights come with tradeoffs. Structural economic models cannot possibly capture every aspect of reality, and any effort to do so would make them unwieldy for either theoretical insight or applied analysis. There will always be some economic choices left out of any particular model—the key question is how to judge what aspects to leave out without rendering the quantitative conclusions of the model irrelevant. The principle we advocate to focus on the question of interest, to achieve parsimony, and to understand how much the model distorts reality is the concept of separability (related to Fisher's separation theorem and Gorman's notion of separability, as discussed in Gorman 1995). This leads to the concept of sufficient statistics, which summarize decisions made outside the model. A specific example is that of consumer two-stage budgeting: the first stage defines the total amount to be spent in a particular period, while the second stage allocates that expenditure to a variety of goods within the period. In modeling the within-period allocation, we may not concern ourselves with what determines the intertemporal allocation problem: under suitable separability assumptions, a sufficient statistic for the intertemporal allocation decision is total consumption (MaCurdy 1983; Altonji 1986; Blundell and Walker 1986; Arellano and Meghir 1992).

The validity of the abstraction of a structural model depends on how appropriate the particular separability assumptions being made are. This sort of abstraction is present even if we are modeling a market equilibrium that considers both the supply and demand sides. We focus on a limited system, because anything more would be too complicated to offer insights. For example, we may model the equilibrium in the labor market and pay determination but say nothing explicitly about the product market or capital investment (for example, Burdett and Mortensen 1998).

Structural economic models should be taught and used as part of the standard toolkit for empirical economists. Of course, other parts of that toolkit include treatment effect models based on quasi-experimental methods and randomized experiments, but these present trade-offs of their own: in particular, the

interpretation of data can become limited and fragmented without the organizing discipline of economic models. Further, without the ability to simulate counterfactuals and more generally to make claims of external validity, the role of empirical analysis is limited to analyzing historical past events without being able to use this accumulated knowledge in a constructive and organized way.

Solving structural models, especially dynamic stochastic models, involves numerical methods. These numerical methods are used to simulate outcomes and counterfactuals as well as to generate moments for use in estimation. The greatest “entry cost” for a researcher wishing to estimate dynamic structural models is learning to solve them, and as we discuss, there are many steps involved in their solution and estimation. Understanding their solution also helps in understanding how they are identified by the data.

In what follows, we start by defining structural models, distinguishing between those that are fully specified and those that are partially specified. We contrast the treatment effects approach with structural models, using Low, Meghir, and Pistaferri (2010) as an example of how a structural model is specified and the particular choices that were made. We follow with a discussion of combining structural estimation with randomized experiments. We then turn to numerical techniques for solving dynamic stochastic models that are often used in structural estimation, again using Low, Meghir, and Pistaferri as an example. The penultimate section focuses on issues of estimation using the method of moments. The last section concludes.

Defining a Structural Model

We begin by differentiating between fully and partially specified structural models, and then consider their relationship to treatment effect models.

Fully Specified Structural Models

Fully specified structural models make explicit assumptions about the economic actors’ objectives and their economic environment and information set, as well as specifying which choices are being made within the model. We call these models fully specified because they allow a complete solution to the individual’s optimization problem as a function of the current information set. In the context of labor economics, Keane and Wolpin (1997) and numerous papers by these authors are prime examples of fully specified structural models. Structural models are the foundation for empirical work in industrial organization with key references being Berry, Levinsohn, and Pakes (1995) and Koujianou-Goldberg (1995); however, most of our discussion draws from examples in labor economics and public finance.

A fully specified dynamic model of consumption and labor force participation will account for how employment and savings decisions are made, taking into account future expectations as well as future implications of these decisions. Working today can imply changes in future wages because of skill accumulation, thus altering the future returns to work and/or through changes in the preferences

for work (habit formation). The choices that the individual makes depend on beliefs about future opportunities (such as wage rates) and future preferences. Thus, in a fully specified model we need to define the distribution of random events (such as shocks to wages and human capital) often specifying the explicit functional form of the distributions and their persistence. We specify the dynamics of other observable or unobservable variables that affect decisions, distinguishing endogenous changes (such as to wealth due to saving decisions, or to human capital as a result of experience) from exogenous changes (such as to prices or to health). These features are all assumed to be in the individual's information set.

Of course no model is literally complete—all models necessarily abstract from possibly relevant choices. These simplifications take two forms: a choice variable may be completely absent from a model, as for example, in the simplest life-cycle model of consumption under uncertainty, which ignores labor supply and takes income to be some exogenous stochastic process. Low (2005) shows that this assumption can lead to underestimates of precautionary saving behaviour. Alternatively we may condition on a choice, but take it as economically exogenous, as discussed in Browning and Meghir (1991). For example, life-cycle behavior may depend on education, but the level of education is taken as given in modeling consumption: the solution of the consumption function will be conditional on education choice.

To illustrate the issues, consider the structural model in Low, Meghir, and Pistaferri (2010). This is a life-cycle model of consumption and labor supply with a specific focus on quantifying employment and wage risk and measuring the welfare cost of risk, with implications for the design of welfare programs. Individuals choose whether to work, whether to change jobs if the opportunity arises, and how much to save.

The first step is to specify the components of the model. A first component is the intertemporal utility function describing preferences and defining what is chosen. A second component is the intertemporal budget constraint, which depends on the available welfare benefits and taxes. Finally, we need to specify how the individual forms expectations about the future, including shocks to human capital and job loss probabilities and opportunities for new jobs. More broadly, we need to specify how preferences are defined over time and over the states of the world and whether the individual is an expected utility maximizer. Together this characterizes the problem facing the individual. These components also define parameters that need to be estimated from the data after we have argued for how they are identified.

We also need to decide what *not* to model. Of course, this list of omissions is a long one, but for models of life-cycle behavior, the most glaring omissions are marriage and fertility: in our example, male preferences are assumed to be separable from these, as is often done in the literature on male labor supply. Education is taken as given (although it affects choices and opportunity sets). Overall savings are explained but not portfolio allocations. Finally, the model is partial equilibrium, in the sense that counterfactual simulations abstract from changes in wages that may result from aggregate changes in the supply of labor. Perhaps more importantly, the model abstracts from aggregate shocks. This means that the results have

little to say about how the welfare effects of idiosyncratic risk vary with the state of the aggregate economy. The judgment is that these other aspects obscure and complicate the model rather than offer important insights given the stated aims. The complications of these extensions are also partly numerical, as we discuss later in this paper.

Some assumptions are made for simplicity and focus, but others are identifying assumptions. For example, the specific distribution of the shocks may be an identifying assumption. A crucial question that arises is the minimal set of assumptions needed for the model to have empirical content and thus be empirically identified. These issues have been much discussed from different perspectives: useful starting points include Rust (1992) and Magnac and Thesmar (2002). Overall, their conclusion is that dynamic discrete choice models need some strong identification assumptions to work. These assumptions can be relaxed somewhat if a continuous outcome variable is involved such as wages (Heckman and Navarro 2007).

The payoff of such assumptions is that we are able to construct a model that is complex in the important dimensions and relatively transparent in the implied mechanisms. In Low, Meghir, and Pistaferri (2010), there are two separate sources of risk—employment and productivity—and a particularly complex budget constraint specifying the details of the available welfare programs. The relative simplicity of the specification hides important numerical complexities because the consumption function may be discontinuous in assets due to the discrete labor supply. The stochastic process of wages is serially correlated, increasing the numerical complexity of the problem. However, within this structure, it is still relatively easy to understand the role of the various sources of risk and how they affect welfare and the way we evaluate various welfare programs. Whether the channel of changed fertility decisions resulting from welfare reform is important for this problem is of course an open question.

Fully specified structural models are particularly useful when we want to understand long-term effects of policy. In a recent paper, Blundell, Costa Dias, Meghir, and Shaw (2016) consider the impact on female careers of tax credits targeted to low-income families with children. A key question is whether tax credits improve longer-term labor market attachment of single mothers by incentivizing them to remain in work and thus avoiding human capital depreciation during the child-rearing period of life. The model quite decisively concludes this is not the case, partly because tax credits in the UK promote part-time work, which is not conducive to building up human capital, and partly because of tax-credit-induced disincentives to work for women within relationships (relative to the situation for single/divorced women). On the other hand, the model also shows that tax credits are by far superior to other commonly used methods of social insurance because of reduced moral hazard. Again, the specification of this model has made a number of simplifying assumptions, the most pertinent of which is to condition on the fertility process and not allow it to change as a result of welfare reform. Despite these sorts of limitations, a structural model that fully specifies behavior can go much further than simply estimating a parameter of interest or testing a

particular theoretical hypothesis. To achieve this, a number of simplifying assumptions have to be made, to maintain feasibility and some level of transparency. The key is that the assumptions are made explicit, allowing future research to question results and make progress on that basis.

The discussion would be incomplete without touching upon empirical equilibrium models. Indeed, there are no better examples of completely specified models than those that also address equilibrium issues, since counterfactual analysis takes into account how the interaction between agents on both sides of the market leads to a new outcome. This requires specifying the behavior of all relevant agents and defining equilibrium in the specific context. At the same time, this provides an excellent example of how studies focus on some key features of equilibrium but not on others; this is both because of the need for focus on a particular question and for keeping modeling and computational complexity in check. Heckman, Lochner, and Taber (1998) and Lee and Wolpin (2008) focus on changes in equilibrium in the labor market; Abbott, Gallipoli, Meghir, and Violante (2013) also focus on the labor market equilibrium but in addition endogenize intergenerational links. Chiappori, Costa Dias, and Meghir (forthcoming), on the other hand, focus on equilibrium in the marriage market and on intrahousehold allocations, but do not consider changes in the labor market equilibrium, keeping wages constant. The search literature focuses on how equilibrium in frictional labor markets affects wage determination, as in the seminal paper of Burdett and Mortensen (1998) and a list of further important contributions too long to discuss here. All these studies estimate equilibrium models in some dimension but abstract from adjustments that are not the prime focus of the study. In so doing, they offer empirical insights on some of the important mechanisms at work in the longer run.

Partially Specified Structural Models

Sometimes our focus is on one component of a fully specified model. Consider an individual who maximizes lifetime utility by choosing consumption, savings, and how much to work in each period. We can derive a within-period labor supply function that is consistent with intertemporal choices but does not fully characterize them. Essentially, this is a reorganization of the marginal rate of substitution between consumption and labor supply. Such models rely on a *sufficient statistic* that summarizes choices not being modeled explicitly. In this case, the sufficient statistic is the amount of consumption allocated to the period. The econometric model defines a relationship between labor supply and wages, conditional on consumption and “looks” like a traditional labor supply model. The model is partially specified, in the sense that there is not enough information to solve for the optimal choice as a function of the information set: for example, the labor supply model resulting from the marginal rate of substitution characterization is silent about expectations for the future, the distribution of shocks, and the functioning of credit markets. However, conditioning on consumption makes the relationship between labor supply and wages valid and dependent upon structural parameters that characterize some aspects of utility. By studying this relationship, we can learn something about

preferences and about the validity of this marginal rate of substitution representation, but we cannot simulate counterfactuals.

This idea builds on the concept of separability and two-stage budgeting introduced by Gorman (for example, Gorman 1995). In the context of empirical labor supply, this approach has been developed by MaCurdy (1983), Altonji (1986), and Blundell and Walker (1986), where separability is a restriction on preferences. More generally, separability is a way of specifying conditions on preferences and technologies that allow us to focus on some aspect of economic behavior without having to deal explicitly with the broader complications of understanding all aspects of behavior at once. In other words, it formalizes what we mean by a partially specified model and offers a way of understanding where misspecification may occur, which would be a failure of the explicit or implicit separability assumptions.

Partially specified structural models are an important empirical tool. They define testable implications for theory and allow us to estimate important parameters (such as the intertemporal elasticity of substitution or the Marshallian wage elasticity) in a way that is robust to different specifications in the parts of the model that remain unspecified, as discussed in the early simultaneous equations literature as well as Browning and Meghir (1991) and recently in Attanasio, Levell, Low, and Sanchez-Marcos (2017), amongst many others. They are explicit about what is kept constant when considering changes in variables and as such can provide consistent estimates for the parameters, given appropriate econometric methods. However, unlike fully specified models, the counterfactual analysis based on these is incomplete: for example, simulating the effect of taxes using a labor supply model that conditions on consumption will be limited by the inability of the model to capture the resulting intertemporal reallocation of consumption.

One of the most analyzed partially specified models is the Euler equation for consumption. It results from an assumption of intertemporally optimizing individuals and rational expectations. It does not require explicit information on the budget constraint because the level of consumption is used as a sufficient statistic for the marginal utility of wealth. This formulation has been the workhorse for examining the presence of liquidity constraints and for estimating the parameter of intertemporal substitution (for example, Attanasio and Weber 1995; Blundell, Browning, and Meghir 1994; Zeldes 1989). The often-used value for the elasticity of intertemporal substitution of about one originates from this body of work. Similarly, much has been learned by the analysis of the Euler equation for investment with adjustment costs (Bond and Meghir 1994). However, for counterfactual analysis, such as the impact of taxation on savings, the model needs to be completed by specifying the full economic environment as discussed above.

Treatment Effect Models

A treatment effect model focuses on identifying a specific causal effect of a policy or intervention while attempting to say the least possible about the theoretical context. The question is: following a policy change (like the introduction of an education subsidy, or a change in a welfare program), can we estimate the

impact on a specific outcome such as education, labor supply, or perhaps transfers between individuals, without specifying a complete model or tying the result to a specific theory? Treatment effect models and their role in program evaluation are developed in Heckman and Robb (1985), Heckman, LaLonde, and Smith (1999), and a subsequent large literature.

The cleanest way of estimating program or treatment effects is experiments where interventions are randomly allocated. Given that in social contexts, compliance with the treatment protocol cannot usually be enforced—that is, subjects allocated to treatment (such as job training) cannot be forced to accept treatment—the randomized experiment will identify the effect of being *offered* treatment, or the intention to treat. Since impacts are possibly heterogeneous, the effect will be an average impact over the population for which randomization took place.

In a treatment model, identification does depend on the assumption that the experiment has not been compromised (either by nonadherence to the protocol or by attrition) and on there being no spillovers from the treatment units (individuals or communities or other groups such as schools) to the control ones, whether directly or through equilibrium mechanisms like price changes and peer effects. Given these important qualifications, we need not assume much about the underlying model of behavior. To get anything more than that out of the experiment, broadening its external validity, will typically require an explicit model, incorporating behavioral and often functional form assumptions.

Sometimes the result of the intention to treat is exactly what we want. However, consider estimating the effect of a welfare program by randomizing its availability (such as randomizing a conditional cash transfer that incentivizes child education and maternal health care, as in Mexico's PROGRESA). The welfare program may change current incentives to work or obtain education, future opportunities, the amount of risk households face, as well as the possibilities of risk-sharing in the communities (for example, Angelucci and De Giorgi 2009). It may even change wages in the affected communities (which it did). The treatment-effects model will isolate the impact of the program on any outcome we look at, but on its own will not be informative about the mechanisms. This limits the lessons from a particular experiment that are generalizable. To obtain more, we need to combine the information from the experiment with a model of household behavior and study how equilibrium in the communities changes. And of course generalizing the results to a scaled-up version of the policy is impossible without a model.

The literature on the effects of taxing higher incomes, discussed in Goolsbee, Hall, and Katz (1999) and Gruber and Saez (2002), provides another example of the issues that arise. Feldstein (1995) measured the impact of decreasing the top tax rate on earnings and incomes by using the 1986 Tax Reform Act. Separate from the issue of the particular merits of this empirical approach to identifying the causal impact, the external validity of the exercise is limited by the fact that the overall effect of reducing the top tax rate depends on how the entire tax schedule was changed and how people are distributed across it, which reduces the generality of the result to the specific context. Even when there is apparent randomization,

such as in the comparison between lottery winners and losers in Imbens, Rubin, and Sacerdote (2001), there is still a threat to external validity: those choosing to participate in the lottery are likely to be those whose behavior will be most affected by winning, as shown by Crossley, Low, and Smith (2016).

Not all treatment effect models are created equal: it is important to distinguish those estimated through randomized experiments from those estimated through quasi-experimental methods, such as difference-in-differences, regression discontinuity, matching, and others. The point of randomized experiments is that results do not depend on strong assumptions about individual behavior, if we are able to exclude the important issues discussed above. However, this clarity is lost with quasi-experimental approaches such as difference-in-differences, where the validity of the results typically depends on assumptions relating to the underlying economic behavior that are left unspecified. For example, Athey and Imbens (2006) show that the assumption underlying difference-in-differences is that the outcome variable in the absence of treatment is generated by a model that is (strictly) monotonic in the unobservables, and the distribution of such unobservables must be invariant over time. These assumptions restrict the class of behavioral models that are consistent with the causal interpretation of difference in differences.

For example, suppose we want to estimate the effects of an intervention to increase the years of education. The difference-in-differences approach assumes that the level of education (in the absence of intervention) will be a strictly monotonic function of just one unobservable. Education is typically driven by the comparison of the benefits of education and the costs of education. The benefit can be expressed as the life-cycle value of wages and other outcomes resulting from an education choice. This benefit will in general be a nonlinear function of heterogeneity in wage returns, particularly if individuals are risk-averse. The costs are also likely to be heterogeneous. So the education choice will generally depend on at least two unobserved components, which are unlikely to collapse into one element of heterogeneity. In this case, the model of education will not satisfy the Athey and Imbens (2006) assumptions and a difference-in-differences analysis of an intervention will not have a causal interpretation.

To make things worse, if the outcome variable is discrete (such as “working” or “not working”) then a point estimate in difference-in-differences can only be achieved by assuming a functional form: the literature is replete with linear probability models estimating impacts using difference-in-differences. These models look simple and straightforward, but the interpretation of their results as causal impacts rely on strong behavioral and functional-form assumptions. In contrast, results from randomized evaluations “only” rely on the integrity of the experiment itself, including of course, the absence of spillovers.

A further issue is the local nature of the results when the impacts of a policy are heterogeneous. This is best illustrated by the regression discontinuity approach, which identifies impacts for individuals who happen to be located close to the discontinuity. Thus while regression discontinuity has some qualities of a randomized experiment (in the sense that being on either side of the discontinuity is

assumed effectively random), in contrast to the experimental approach, the impact is local to a very specific group of people defined by proximity to the discontinuity. These concerns are more broadly relevant for quasi-experimental approaches as discussed in Imbens and Angrist (1994) and in Heckman and Vytlačil (2005).

In short, randomized experiments provide causal effects without having to refer to a specific economic model or structure. Quasi-experimental approaches on the other hand, while not focusing on structural parameters, rely on underlying assumptions about behavior that potentially limit the interpretability of the results as causal. The attraction of these approaches is their simplicity. However their usefulness is limited by the lack of a framework that can justify external validity, which in general requires a more complete specification of the economic model that will allow the mechanisms to be analyzed and the conclusions to be transferred to a different set of circumstances. This is one of the key advantages of structural models: they describe the mechanisms through which effects operate and thus provide the framework for understanding how a particular policy may translate in different environments.

Combining Randomized Experiments and Structural Modeling

A combination of a fully specified model and randomized experiments can enhance analysis in ways that either of the two approaches alone would miss. Indeed, one of the most important recent advances in empirical economics uses dynamic structural models with exogenous sources of variation. The idea is of course not new and goes back at least to Orcutt and Orcutt (1968) as well as to the evaluation of the Gary negative income tax experiment in Burtless and Hausman (1978). Also, Rosenzweig and Wolpin (1980) combine information from quasi-experimental variation to infer structural relations in a twins study to analyze the quality–quantity model of fertility.

The renewed interest in this approach brings together the advantages of credible evaluation that relies on randomized experiments or (arguably) exogenous variation induced by policy changes, with the systematic economic analysis of structural models. A couple of prominent examples include Blundell, Duncan, and Meghir (1998) who use changes in the structure of wages and tax policy reforms to identify a partial model of labor supply, and Kaboski and Townsend (2011) who estimate a model of household investment and borrowing and validate its predictions using Thai data drawn from an expansion of microfinance availability in a large set of villages.

Experimental evidence can be used either to validate a structural model or to aid in the estimation process. These two alternative ways of using the same experimental evidence can be illustrated by comparing Todd and Wolpin (2006) and Attanasio, Meghir, and Santiago (2012). In 1998, the Mexican government experimented with a conditional cash transfer program whose intention was to increase the school participation of children in poor rural areas and improve preventive

health care participation by mothers. PROGRESA, as the program became known, was to be evaluated by a cluster randomized control trial. Out of a population of 506 poor rural communities, 320 were assigned to receive the program immediately, while the remaining ones were kept back as a control, only receiving the program two years later. PROGRESA consisted of offering nutritional supplements to young children and a subsidy to families (disbursed to the mother) conditional on children's attendance at school. Mothers had to attend health clinics regularly to be eligible.

The intervention was highly successful. Schultz (2004) carried out the main evaluation of the program and shows that schooling participation increased. But can we learn more from the experiment and the associated data than the magnitude of the treatment effect? Specifically, can we say something about the design of the program and more generally, something about how costs of schooling affect educational participation?

In a standard economic model, the conditional school grants change school participation by counteracting the opportunity cost of schooling. Todd and Wolpin (2006) use this insight to validate a model of education attendance and fertility. They estimate the model based on data from the control group only. They then predict the impact of the experiment by reducing the wage in the model by an amount equivalent to the grant when the child went to school. They thus use the experiment to validate a dynamic model, which as specified, is identifiable from the control data only.

Using data from the experiment, Attanasio, Meghir, and Santiago (2012) identify a richer model (in some dimensions) that implies a more general cost-of-school function. Like Todd and Wolpin (2006), they set up a forward-looking model of educational choice through high school, where the individuals and their families decide each period whether to attend school. The benefits of schooling accrue in the future through better labor market opportunities, identified by the observed schooling attendance in the control group. A more general specification would use observations on the subsequent career to improve identification. The cost of schooling is affected by four elements: 1) forgone child labor income; 2) the amount of the PROGRESA grant for which the individual is eligible if attending school, which varies by the age of the child; 3) past school attendance, which may reduce cost because of habits or because past learning makes schooling now easier; and 4) an unobserved cost of attending school associated with the child's scholastic ability.

This structural model is explicitly dynamic: each year of schooling adds to human capital and future standards of living; there is uncertainty over whether the child will pass the grade; the grant is only available up until age 18; and current attendance affects future costs. The key point is that the model in Attanasio, Meghir, and Santiago (2012) allows the PROGRESA grant to have a different effect from the wage: the authors use the experimental variation to identify this extra effect. The finding is that the impact of the grant is larger than that predicted by the changes in school attendance as a function of forgone wages. This finding poses important

questions of interpretation, but it highlights that the experiment allows the model to be extended and to address directly whether the grant has a different effect from the standard opportunity cost. The use of the experiment allowed the relaxation of some of the restrictions from economic theory, thus broadening the scope of the model and the interpretation of the experimental results. A further development of the model would require explaining why the grant has a different impact than forgone wages. Attanasio, Meghir, and Santiago speculate that this has to do with intrahousehold allocations and the fact that the grant goes to women. From the point of view of the discussion here, progress in understanding would require adding such an intrahousehold component and thinking about ways to identify it.

Of course, it is important not to overstate the synergies between structural models and experiments: in most cases, randomized experiments only offer discrete sources of variation—policy is on or off—which is far from the requirements for identification in dynamic models, which would typically require continuous variation (Heckman and Navarro 2007).

The above example illustrates how the experiment can add to the identification potential of the structural model. But what does the structural model add to the experiment? We know how the experiment affected school attendance at various ages. What does the model offer in addition to this finding, and what are the assumptions on which any additional insights are based?

The basic gain from using the structural models is that they allow a better understanding of the mechanisms and analysis of counterfactuals. Attanasio, Meghir, and Santiago (2012) focus on counterfactuals: for example, we can ask whether the grant, which varies by age of the child, would be more effective if structured differently over age, holding total financial cost constant. In terms of mechanisms, Attanasio, Meghir, and Santiago discuss the potential role of intrahousehold allocations; but the age limitation of the grant is an important factor in its effect. They also estimate whether the impact on wages resulting from the change in child labor supply dampened significantly the effect of the program—it did not. A richer model could look at how the program affected risk and risk sharing in the community, thus changing decisions including that of school attendance. Todd and Wolpin (2006) also investigate impacts on fertility. These rich behavioral models offer a deeper insight of just how an intervention can affect the final outcome. Understanding the mechanisms is central to designing policies and avoiding unintended effects as well as for building a better understanding of whether a policy can reasonably be expected to work at all.

The extra richness offered by the model does not come for free. We need to make additional assumptions that were not required for the simple experimental evaluation. Consider the counterfactual that restructures the PROGRESA grant by adjusting the amounts offered at each age. The ability to assess this proposal depends on knowing how education participation varies with the grant at different ages. The amount of the grant did vary with age; however, each age is associated with just one amount mandated by the program—there is no age-by-age experimental variation (although conceptually there could be). To recover how the

effect of the grant varies by age, we need to assume that this effect varies smoothly and does not follow the exact pattern of variation of the grant by age. One can be understandably skeptical of results that rely on untestable assumptions about preferences. However, the assumptions in these models tend to be explicit, so promoting transparency and allowing for explicit criticism and improvements. For the purpose of this paper, it provides a good example of the types of assumptions that often need to be made to extend the narrow conclusions of an experiment to a broader context. In a more complex experiment, one can imagine that the amounts themselves within the experiment would be randomized at each age—thus offering stronger identification of this effect. In practice, it is very hard to implement experiments that are complex enough to offer variation in all the directions required for identification of all desired insights and still have sample sizes to allow sufficient statistical power.

Structural models include further restrictions. For example, they often require assumptions about the distribution of random preferences. In Attanasio, Meghir, and Santiago (2012), it is assumed that psychic costs of education can be described by a mixed logistic distribution. A central question in this literature is whether such assumptions are needed, or if they could be relaxed with richer data. In an enlightening paper, Magnac and Thesmar (2002) argue convincingly that a dynamic discrete choice model, such as the one in Attanasio, Meghir, and Santiago (2012), does depend on identifying assumptions relating to the distribution of preferences. The reason is quite intuitive because all outcome variables are discrete.

More can be achieved in models with continuous variables: Heckman and Navarro (2007) develop identification in a dynamic Roy model of education and wages. As they emphasize, the key to identification of the dynamic model is that they use information on measured consequences of treatment—for example, on wages. They show that identification restrictions can be relaxed if one observes explicitly a continuous outcome variable, such as the wage rate, and if the dynamic discrete choice depends on some continuous variable with large support, such as school fees (see also Meghir and Rivkin 2011). In practice, such conditions are usually not met and the functional form restrictions will play a role in analyzing the actual dataset. Heckman and Navarro (2007) also emphasize the use of cross-equation restrictions implied by the theoretical structure of the model. Here there is an important distinction to be made between restrictions on the shape of distributions of unobservables, for which there is rarely any theory, and restrictions that follow a clear reasoning and foundation in theory. While we should minimize ad hoc restrictions, it is also important to realize that empirical analysis can never do away with theoretical foundations and still remain useful as a learning tool. Data from a randomized experiment can however be very helpful here, either in showing that despite the various functional form assumptions, the model matches the unbiased results of the experiment, or in using the experiment to ensure that the resulting estimates reproduce the impacts. In this sense, the experiment can aid in identification.

Combining randomized experiments and credible quasi-experimental variation with structural models seems to bring together the best of both approaches of empirical economics: it identifies causal effects non- or semi-parametrically for specific policies, provides useful identifying information for the structural model, and offers a coherent way for understanding mechanisms and counterfactuals through the organizing lens of economic theory. This approach is growing in influence: beyond the papers already cited, Duflo, Hanna, and Ryan (2012) use a structural model to analyze the results of a school monitoring experiment in India; Kaboski and Townsend (2011) combine information from quasi-experimental evidence from the Thailand Million Baht Village Program with a structural model of small family businesses to understand the mechanisms underlying the workings of microfinance (see also Garlick 2016); and Voena (2015) uses differences-in-differences to evaluate the effect of divorce laws on household behavior and then uses this data to fit a dynamic intrahousehold model with limited commitment in order to analyze policy counterfactuals.

Solving Structural Models

The specification issues discussed above are driven both by the importance of a well-focused and empirically identified economic model as well as by computational feasibility. There have been huge advances in both computational methods and power over the last 30 years allowing much more flexibility in what can be implemented in practice. However, computational constraints remain and to some extent will always be with us. In this section, we discuss computational issues relating to solving these models, which is where most of the difficulty lies. We use Low, Meghir, and Pistaferri (2010) loosely as a case study and discuss in particular the computational implications of relaxing the separability assumptions.

In some situations, structural estimation is simple and relies on linear methods: for example, estimating demand systems in static models or estimating Euler equations with complete markets as in Altug and Miller (1990) or even with incomplete ones as in Meghir and Weber (1996). But more often than not, structural models and particularly dynamic stochastic models involve nonlinear estimation, and require numerical methods to solve the model to generate moments for estimation. The greatest “entry cost” for a researcher wishing to estimate dynamic structural models is learning to solve such models accurately and efficiently. For a broader textbook discussion of these methods, useful starting points include Adda and Cooper (2003) and Miranda and Fackler (2002). Some more recent, faster methods are due to Carroll (2006), Fella (2014), and Barillas and Fernández-Villaverde (2007).

The heart of solving dynamic structural models is the computation of value functions and corresponding decision rules. The value functions associate a numerical value to a decision or set of decisions, conditional on the relevant state variables, and conditional on all future decisions being optimal. The decision rules describe how individuals behave following different realizations of the economic

environment. The state variables describe the economic environment. These include variables that are independent of past choices by the agent (and so are treated as exogenous) as well as variables that evolve depending on past decisions (and so are endogenous).

The discussion below relates to finite horizon life-cycle models. In these models, age is part of the state space—which means that the value functions are not stationary. There is a class of structural dynamic models with infinite horizons in which the value functions are stationary, not depending on age. Equilibrium search models of the labor market are usually specified in this way for purposes of convenience. The solution methods for these are related but different, and not touched upon here.

In Low, Meghir, and Pistaferri (2010), the decision rules describe whether an individual at each age would choose to work and how much the individual would save. Decision rules are obtained by comparing the value functions derived from different choices, at a given state of the economic environment. The state variables are wealth, individual productivity, and the matched firm type. The model makes numerous separability assumptions, as discussed above, especially over fertility and marriage: neither children nor marriage are considered choices in the model, and preferences over consumption and employment are assumed to be separable from marriage and fertility. Relaxing these assumptions expands the choice set, increasing the number of value function comparisons. It would also expand the state space to include current marital status, details of any partner, as well as family size, increasing the description of the economy and increasing the number of points at which decision rules have to be solved. The value of the separability assumptions therefore is in reducing the computational burden, as well as in making the model less opaque.

Armed with these decision rules, the researcher can then simulate behavior. There are four core steps in solving a dynamic structural economic model.

In the first step, the points (or nodes) of the state space at which the model needs to be solved are specified. Numerical solution requires defining the bounds of each of the variables, so that one can then think of a multidimensional grid of state variables. The state space fully specifies all aspects of the economic environment which affects the particular choices being analyzed in the model. In Low, Meghir, and Pistaferri (2010), the grid of specific values of the exogenous state variables, such as the permanent wage, are set before solving the model, using approximations to transition probabilities as in Adda and Cooper (2003, chapter 3). For endogenous state variables, such as with wealth, restricting the number of values to a discrete set would restrict the choice set in an arbitrary way. In Low, Meghir, and Pistaferri, this would have meant a discrete set of consumption values, which would introduce jumps in behavior that are not observed in the data. We keep these endogenous state variables continuous. Discretization can nonetheless be used to determine the points where the model is actually solved and then interpolation can be used between these points.

There are many alternative ways to interpolate: linear interpolation is usually robust, particularly when decision rules are not smooth, and in Low, Meghir, and

Pistaferri (2010), we start with this approach. Other methods of interpolation include imposing assumptions about smoothness in the decision rules, which then allows fitting either higher-order local splines or, alternatively, polynomials across the whole state space to provide a global approximation. The tradeoff here is that linear interpolation tends to require more points spanning the state space. Alternative methods need fewer points in the state space but impose further restrictions on the form of the solution.

Carroll (2006) and Fella (2014) discuss using endogenous grid points which are relevant for endogenous state variables. Getting the minimum value of the grid right requires care: in Low, Meghir, and Pistaferri (2010), the lower bound on assets is determined by an exogenously set borrowing limit. Alternatively, the lower bound can be determined by a “no-bankruptcy condition,” specifying that borrowing has to be limited to what can be repaid with certainty—a “natural” liquidity constraint.

In the second step, we specify a terminal condition defining the continuation value in the subsequent periods beyond which we do not model decisions. In Low, Meghir, and Pistaferri (2010), the terminal condition is death, but it does not have to be: in Attanasio, Meghir, and Santiago (2012), the terminal condition is defined by the oldest age the child could attend high school, taken to be 18. In general the terminal value is a function of the state variables at that point. In Attanasio, Meghir, and Santiago (2012), the state variable is whatever schooling the child has accumulated by that age. The structure of this terminal value function is either tied directly to the model, with no new parameters, or it needs to be estimated with the rest of the model. Choosing an appropriate terminal point consistent with the model can economize on parameters to be estimated and improve identification. For example, if it is reasonable to assume that no individual lives beyond say 110 and that there are no bequests other than accidental ones, then the terminal value is defined explicitly by the problem and no extra arbitrary modeling assumptions have to be imposed.

In the third step, the “value function” at each node in the state space specified on the grid is solved, starting with the terminal period. The solution involves solving a numerical optimization problem, or a nonlinear equation solution to a first-order condition. The model in Low, Meghir, and Pistaferri (2010) contains a mixture of discrete and continuous choices: over whether to participate in the labor market, and over how much to save. This combination of discrete and continuous choices raises a problem, because changes in asset holdings can lead to changes in participation status and to jumps in the decision rule for consumption. We deal with this nonconcavity by solving for value functions conditional on the discrete choice, and then taking the maximum over these. The number of conditional value functions to solve increases with the number of discrete choices. However, these conditional value functions may themselves not be concave. The solution, numerically, is that if there is “enough” uncertainty about changes in future prices or wages then the expected value function will typically be concave. Nevertheless, this can rarely be proved and depends on the amount of uncertainty in the model. In practice, this means that we need to investigate numerically whether multiple solutions occur.

The fourth and final solution step is to iterate backwards one period at a time, at each period solving for each point in the state space. The solution in earlier periods will be determined taking account of expectations about future outcomes (based on the distribution of potential shocks) and also how the individual will respond in the future to those outcomes (based on the already-solved future decision rules. Because expectations have to be calculated, this involves numerical integration over the unknown random variables: for example, in Low, Meghir, and Pistaferri (2010), these are shocks to wages, job offer arrivals, and firm types. The more underlying random variables are involved, the higher the dimension of integration and consequently the computational costs can rise exponentially. This factor limits in practice the amount and source of uncertainty that one can introduce in a model. Notice also that the distribution of the shocks may depend on past realizations (rather than being independent and identically distributed). For example, if shocks to wages are serially correlated, the realization of a future shock will depend on the value of the current shock. This means that the current shock to wages is in the state space: an extra exogenous continuous state. For this reason, the way we specify the distribution of random events is very important in keeping the problem tractable.

The decision rules solved in this way by backward induction specify what an economic actor will choose given any particular realization of the state of the world. These decision rules will then be combined with particular randomized realizations of the stochastic variables starting at an initial period and simulating forwards. In Low, Meghir, and Pistaferri (2010), the randomized realizations are of permanent shocks to wages, and of wage offers, firm type, and job destruction. These stochastic shocks are responsible for life-time career paths being so different for what otherwise appear to be identical individuals. Inputting one complete set of realizations of these stochastic variables into the decision rules generates the life-cycle path for consumption and labor supply for one individual. This calculation is then repeated a number of times to generate average life-cycle profiles, along with other moments that are needed. We return below to the issue of the number of repetitions when discussing implementation of this approach using the simulated method of moments.

Using Method of Moments for Estimation of Structural Models

The numerical solution is used to generate predictions about behavior for a given set of parameter values. These parameter values need to be estimated.

Estimation of dynamic structural models involves nonlinear optimization with respect to the unknown parameters. However, the key difficulty with this estimation is that we cannot express analytically the functional relationship between the dependent variables and the unknown parameters. In order to see how a change in a parameter changes the dependent variable, the entire model solution needs to be generated afresh. If solving the model once already takes time, the problem is compounded by estimation that requires solving the model repeatedly. Moreover, numerical approximation errors in the solution of the model can compound

the estimation complexity. There is an active literature on the way to approach the problem: one is the nested fixed point algorithm (Rust 1987), where the model is solved for each set of parameters that are tried out by the numerical optimization algorithm. A recent alternative, which under certain circumstances can be faster, is the method of Su and Judd (2012).

Beyond the choice of algorithm for optimization, another important choice is the criterion function to be optimized as a function of the parameters. Traditionally, maximum likelihood was used for estimating structural models. This approach is most efficient, exploiting all the information in the specification. However, constructing the likelihood function is impossible or computationally intractable for many models. Estimation now typically uses the method of moments (or indirect inference) (for a formal discussion, see McFadden 1989; Pakes and Pollard 1989; Gourieroux, Monfort, and Renault 1993). For our purposes, we use the term moments in a broad sense to mean any statistic of the data whose counterparts can be computed from model simulations for a given set of model parameters. For example, moments include means, variances, and transition rates between states, as well as regression coefficients from simple “auxiliary” regressions.

With the method of moments, it is easier to tell which features of the data identify which structural parameters. Further, use of multiple datasets is straightforward and the researcher can put emphasis on fitting moments central to the analysis. Finally, this method eliminates the computational burden of using enormous administrative datasets with millions of observations: the data moments need only be computed once; the computational burden will then be due exclusively to the time it takes to solve the model. The downside of this approach is that it does not use all information in the data, and we do not have an easily implementable way of defining which moments need to be used to ensure identification. One must carefully define what are the key features of the data that will identify the parameters. Moreover, in finite samples, the results may be sensitive to the choice of moments.

In Low, Meghir, and Pistaferri (2010), we need to estimate the parameters governing the opportunity set, which include the wage process, job destruction, job arrival rates on and off the job, fixed costs of working, and the parameters, which include the discount rate, elasticity of intertemporal substitution, and disutility of work. Estimation of these parameters can be described in a five-step algorithm:

- 1) Start with an initial guess at a set of parameter values θ .
- 2) Numerically solve the model given the parameter vector θ (as described in the previous section).
- 3) If individuals are *ex ante* identical, simulate the careers of say S individuals using a random number generator for realisations of the stochastic variables, and construct moments from the simulated moments analogous to those constructed from the data. If individuals differ by exogenous observed factors, simulate S careers for each value of the exogenous initial conditions. Similarly if individuals differ by some unobserved characteristic (whose distribution is estimated together with the rest of the model) again simulate S careers for each point of support of the unobservable and then take suitable weighted averages when constructing the moments.

4) Calculate the “criterion function” being minimized. This may be a simple or weighted quadratic distance between the data and the simulated moments.

5) Update the set of parameters θ to minimize the criterion function and return to step 2 and numerically solve the model with the updated parameters.

There are many decisions in implementing this algorithm. Here we discuss the main ones: what parameters to estimate, what moments to use, how to weight the moments, how to optimize to minimize the criterion, and post-estimation, what checks to carry out.

Choosing the Parameters to Estimate

A fully specified economic model requires that all parameters governing the opportunity set and preferences be determined, which can often make the problem unmanageable. The set of parameters can be divided into three: First, some parameters of the economic environment can be obtained directly from the institutional setting or data, requiring the assumption that the particular aspect of the environment is not affected by economic choices made within the model. For example, in Low and Pistaferri (2015), the specification of how health shocks evolve was estimated directly from the data, requiring the assumption that labor supply and other choices did not affect health.

Second, some parameters can be obtained using a partially specified model. Parameters estimated in this way are robust to details of the fully specified structural model. For example, Attanasio, Levell, Low, and Sanchez-Marcos (2017) use an Euler equation to estimate the elasticity of intertemporal substitution to use in a fully specified model. Low, Meghir, and Pistaferri (2010) and Low and Pistaferri (2015) estimate the wage process using a reduced-form procedure with the residuals identifying the wage uncertainty. The disadvantage of the procedure is that estimation is not completely in tune with the fully specified model. However, what may seem to be a shortcoming can also be an advantage: using the partially specified model means many auxiliary assumptions are not imposed on all components of the model.

Finally, parameters that are the key drivers of the economic choices in the model form part of the full structural estimation. In Low, Meghir, and Pistaferri (2010), these parameters were the disutility of work, the fixed cost of work, and job market frictions. In Low and Pistaferri (2015), these parameters also included the acceptance probabilities onto disability insurance.

Selecting Moments

More moments are not necessarily helpful in practice: moments need to be economically important to the model and informative about parameters. In Low, Meghir, and Pistaferri (2010), key moments were employment rates and unemployment duration at different ages. Employment rates were related to fixed cost of work, and durations were related to job arrival rates, although both sets of parameters affect both moments through the structure of the model. Moments used may include reduced form regressions, population means, or elasticities from the

literature. Low and Pistaferri (2015) use coefficients from a regression of consumption on health status as moments to inform how health shocks affect the marginal utility of consumption. Other important moments may be transition rates, dispersion, and the time series properties of wages.

Simulating these moments in step 3 of the algorithm above requires randomly generated variables to represent the exogenous stochastic processes in the structural model. For each individual simulated, there is a random realization for each stochastic process. The complete set of random numbers for all individuals should be generated only once at the start of the estimation, and the same set of random numbers should be used in each iteration of the criterion function. As the number of simulations increases to infinity, the simulation error goes to zero, implying the moments become equal to the theoretically implied ones. At this point we are only left with the usual sampling error from the data. In general, due to the number of simulations being finite, simulation error should be taken into account in computing the standard errors of the estimated parameters.

The distributions of the stochastic processes may depend on parameters that need to be estimated. In order to make sure the underlying random draws are the same across iterations we need to draw uniform $(0, 1)$ random variables that can then be transformed to follow whatever distribution the model implies (for example, $N(0, \sigma^2)$ where σ^2 is estimated).

Weighting the Moments

Moments may not be of equal economic importance, or measured with equal precision, or measured in comparable units. These considerations determine the choice of weighting matrix on the moments. Alternatives are the inverse of the full variance–covariance matrix, the inverse of the diagonals of the variance–covariance matrix, the identity matrix, or conversion of deviations into percentage deviations.

The “optimal weighting matrix” is the inverse of the variance–covariance matrix of the moments. This puts greater weight on more precisely estimated moments, and corrects the weighting on moments that are correlated. Ruge-Marcia (2012) shows the advantages of this weighting in a Monte Carlo exercise. However, with small samples, Altonji and Segal (1996) emphasize that the identity matrix (that is, equal weighting) may be the best choice because using hard-to-estimate higher-order moments of the data that constitute the weight matrix may actually introduce substantial bias. Equal weighting does not differentiate the precision with which each moment is estimated, and the units of measurement affect the weighting. The moments can be normalized to convert the difference between moments into the percentage deviation, which is equivalent to using a matrix of the inverse of the moments in the data squared. An alternative is the inverse of the diagonals of the variance–covariance matrix, but the issue remains that the more precisely measured moments get more weight, regardless of how important the moments are for the question at hand.

In Low, Meghir, and Pistaferri (2010), labor participation rates are precisely measured, whereas duration-of-unemployment numbers are imprecisely measured.

Weighting based on precision with which moments are estimated would have meant durations would fit poorly, reducing the relevance of the model. In that study, we reduce the scope of this problem by using only economically relevant moments of the data and converting the moments to be percentage deviations.

Optimization with Simulated Moments

Simulated moments are often not smooth with respect to the parameters and as a result, derivative-based methods of optimization are often inappropriate. A straightforward method is Dantzig's classic simplex method. The simplex method is derivative-free and while it can be computationally slow, it is robust. Recently, Markov Chain Monte Carlo methods for optimization have become more common. This approach requires no more than simulating the model and computing moments given a set of parameters. Chernozhukov and Hong (2003) have shown how Markov Chain Monte Carlo can provide estimators that are asymptotically equivalent to minimizing the method-of-moments criterion. While the Markov Chain Monte Carlo can be slow to converge on some occasions, in practice other alternatives may be much worse. Many researchers make use of parallel computing with multiple chains running at the same time.

Standard Errors and Post-Estimation Checks

After the parameters are estimated in a structural economic model, the list of tasks is not yet complete: additional checks are needed.

First, calculate parameter standard errors. Various papers on simulated method of moments and indirect inference like McFadden (1989), Pakes and Pollard (1989), and Gourieroux, Monfort, and Renault (1993) provide the appropriate results. A practical difficulty is that these approaches require derivatives of the moments with respect to the parameters of the model. Another difficulty arises from the fact that the estimation error in pre-estimated parameters also needs to be taken into account. This correction can become computationally hard.

Second, use the finite difference approach in the first check to show how moments change with estimated structural parameters. This information helps make estimation more transparent, by showing which parameters are pinned down by which moments.

Third, show the 95 percent confidence interval for the difference between each simulated moment and its data counterpart. This provides a metric for judging how well moments fit. In Low and Pistaferri (2015), for example, the model could not match the participation rates of the healthy who were over 45.

Finally, consistency should be checked between any estimates from the partially specified or pre-estimation stage and the implications of the fully specified model. In Low and Pistaferri (2015), the wage process was pre-estimated with a reduced form selection correction. Data on the simulated individuals was used after the estimation to check consistency of the full model with the selection model. A further test compares simulated predictions with additional moments or reduced form evidence, preferably not targeted in estimation. An important validation of Low

and Pistaferri was to compare the simulated elasticities of the receipt of disability insurance with respect to generosity to the reduced form estimates in the literature.

The ultimate purpose is to produce an estimated model that is internally consistent, so the estimates can be used for counterfactual analysis. Being explicit about each of these steps can help to provide transparency about the mechanisms and the sources of identification.

Conclusion

Structural economic models are at the heart of empirical economic analysis, offering an organizing principle for understanding data, for testing theory, for analyzing mechanisms through which interventions operate, and for simulating counterfactuals. It has been long understood that econometric identification of such models will necessarily depend on prior assumptions and on theory; but without the organizing device of theory, it is impossible to make progress in our understanding. We argue that the resurgence and increased popularity of the idea of combining randomized experiments or plausible quasi-experimental variation together with structural economic models can strengthen the value of empirical work substantially. Indeed, researchers should think more ambitiously and use theory to define experiments that need to be run to test and estimate important models.

Structural economic models are difficult to use because of computational complexity. Moreover, it is easy to end up with overcomplicated and unwieldy models that offer little insight into mechanisms and whose identifiability is, to say the least, obscure. The trade-off between providing the necessary complexity to be economically meaningful and maintaining transparency is at the heart of good structural modeling. Our approach is to be explicit about what separability assumptions can be invoked: a fully specified structural model will not capture all choices, but will be explicit about which choices are part of the model and which choices are not, and will solve explicitly for all choices in the model. Choices can be left out of a model if they do not affect the choices we are modeling, due to separability in preferences.

With the increasing use of structural models and the progress of both computational power and numerical methods, the economics profession is becoming much more familiar and skilled in the specification and use of structural models. In our view, this all for the good, and it is hard to see how progress can be achieved without both sides of empirical work: experiments generating exogenous variation, and theory-based models.

■ *We thank Tom Crossley, Felix Grey, Gordon Hanson, Enrico Moretti, Ann Norman, and Timothy Taylor for excellent comments and suggestions. Costas Meghir has been supported by the Cowles Foundation and the Institution for Social and Policy Studies (ISPS) at Yale University. All errors and interpretations are our own.*

References

- Abbott, Brant, Giovanni Gallipoli, Costas Meghir, and Gianluca L. Violante. 2013. "Education Policy and Intergenerational Transfers in Equilibrium." NBER Working Paper 18782.
- Adda, Jérôme, and Russell W. Cooper. 2003. *Dynamic Economics: Quantitative Methods and Applications*. MIT Press.
- Altonji, Joseph G. 1986. "Intertemporal Substitution in Labor Supply: Evidence from Micro Data." *Journal of Political Economy* 94(3, Part 2): S176–S215.
- Altonji, Joseph G., and Lewis M. Segal. 1996. "Small-Sample Bias in GMM Estimation of Covariance Structures." *Journal of Business and Economic Statistics* 14(3): 353–66.
- Altug, Sumru, and Robert A. Miller. 1990. "Household Choices in Equilibrium." *Econometrica* 58(3): 543–70.
- Angelucci, Manuela, and Giacomo De Giorgi. 2009. "Indirect Effects of an Aid Program: How Do Cash Transfers Affect Ineligibles' Consumption?" *American Economic Review* 99(1): 486–508.
- Arellano, Manuel, and Costas Meghir. 1992. "Female Labour Supply and On-the-Job Search: An Empirical Model Estimated Using Complementary Data Sets." *Review of Economic Studies* 59(3): 537–59.
- Athey, Susan, and Guido W. Imbens. 2006. "Identification and Inference in Nonlinear Difference-In-Differences Models." *Econometrica* 74(2): 431–97.
- Attanasio, Orazio P., Peter Levell, Hamish Low, and Virginia Sanchez-Marcos. 2017. "Aggregating Elasticities: Intensive and Extensive Margins of Female Labour Supply." Cambridge Working Papers in Economics 1711.
- Attanasio, Orazio P., Costas Meghir, and Ana 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.
- Attanasio, Orazio P., and Guglielmo Weber. 1995. "Is Consumption Growth Consistent with Intertemporal Optimization? Evidence from the Consumer Expenditure Survey." *Journal of Political Economy* 103(6): 1121–57.
- Barillas, Francisco, and Jesus Fernández-Villaverde. 2007. "A Generalization of the Endogenous Grid Method." *Journal of Economic Dynamics and Control* 31(8): 2698–2712.
- Berry, Steven, James Levinsohn, and Ariel Pakes. 1995. "Automobile Prices In Market Equilibrium." *Econometrica* 63(4): 841–90.
- Blundell, Richard, Martin Browning, and Costas Meghir. 1994. "Consumer Demand and the Life-Cycle Allocation of Household Expenditures." *Review of Economic Studies* 61(1): 57–80.
- Blundell, Richard, Monica Costa Dias, Costas Meghir, and Jonathan Shaw. 2016. "Female Labor Supply, Human Capital, and Welfare Reform." *Econometrica* 84(5): 1705–53.
- Blundell, Richard, Alan Duncan, and Costas Meghir. 1998. "Estimating Labor Supply Responses Using Tax Reforms." *Econometrica* 66(4): 827–61.
- Blundell, Richard, and Ian Walker. 1986. "A Life-Cycle Consistent Empirical Model of Family Labour Supply Using Cross-Section Data." *Review of Economic Studies* 53(4): 539–58.
- Bond, Stephen, and Costas Meghir. 1994. "Dynamic Investment Models and the Firm's Financial Policy." *Review of Economic Studies* 61(2): 197–222.
- Browning, Martin, and Costas Meghir. 1991. "The Effects of Male and Female Labor Supply on Commodity Demands." *Econometrica* 59(4): 925–51.
- Burdett, Kenneth, and Dale T. Mortensen. 1998. "Wage Differentials, Employer Size, and Unemployment." *International Economic Review* 39(2): 257–73.
- Burtless, Gary, and Jerry A. Hausman. 1978. "The Effect of Taxation on Labor Supply: Evaluating the Gary Negative Income Tax Experiment." *Journal of Political Economy* 86(6): 1103–30.
- Carroll, Christopher D. 2006. "The Method of Endogenous Gridpoints for Solving Dynamic Stochastic Optimization Problems." *Economics Letters* 91(3): 312–20.
- Chernozhukov, Victor, and Han Hong. 2003. "An MCMC Approach to Classical Estimation." *Journal of Econometrics* 115(2): 293–346.
- Chiappori, Pierre-Andre, Monica Costa Dias, and Costas Meghir. Forthcoming. "The Marriage Market, Labor Supply and Education Choice." *Journal of Political Economy*.
- Crossley, Thomas F., Hamish Low, and Sarah Smith. 2016. "Do Consumers Gamble to Convexify?" *Journal of Economic Behavior and Organization* 131: 276–91.
- Duflo, Esther, Rema Hanna, and Stephen P. Ryan. 2012. "Incentives Work: Getting Teachers to Come to School." *American Economic Review* 102(4): 1241–78.
- Feldstein, Martin. 1995. "The Effect of Marginal Tax Rates on Taxable Income: A Panel Study of the 1986 Tax Reform Act." *Journal of Political Economy* 103(3): 551–72.
- Fella, Giulio. 2014. "A Generalized Endogenous Grid Method for Non-smooth and Non-concave Problems." *Review of Economic Dynamics* 17(2):

329–44.

Garlick, Julia. 2016. “Essays in Development Economics.” Yale PhD thesis.

Gorman, W. M. 1995. *Collected Works of W.M. Gorman*. Edited by C. Blackorby and A. F. Shorrocks. Oxford University Press.

Goolsbee, Austan, Robert E. Hall, and Lawrence F. Katz. 1999. “Evidence on the High-Income Laffer Curve from Six Decades of Tax Reform.” *Brookings Papers on Economic Activity* no. 2, pp. 1–64.

Gourieroux, Christian, Alain Monfort, and Eric Renault. 1993. “Indirect Inference.” *Journal of Applied Econometrics* 8(S1): S85–S118.

Gruber, Jon, and Emmanuel Saez. 2002. “The Elasticity of Taxable Income: Evidence and Implications.” *Journal of Public Economics* 84(1): 1–32.

Hansen, Lars P. 1982. “Large Sample Properties of Generalised Method of Moments Estimators.” *Econometrica* 50(4): 1029–54.

Heckman, James J., Robert J. LaLonde, and Jeffrey A. Smith. 1999. “The Economics and Econometrics of Active Labor Market Programs.” Chap. 31 in *Handbook of Labor Economics*, vol. 3, part A, edited by David Card and Orley Ashenfelter, pp. 1865–2097.

Heckman, James J., Lance Lochner, and Christopher Taber. 1998. “Explaining Rising Wage Inequality: Explorations with a Dynamic General Equilibrium Model of Labor Earnings with Heterogeneous Agents.” *Review of Economic Dynamics* 1(1): 1–58.

Heckman, James J., and Salvador Navarro. 2007. “Dynamic Discrete Choice and Dynamic Treatment Effects.” *Journal of Econometrics* 136(2): 341–96.

Heckman, James J., and Richard Robb Jr. 1985. “Alternative Methods for Evaluating the Impact of Interventions: An Overview.” *Journal of Econometrics* 30(1–2): 239–67.

Heckman, James J., and Edward Vytlacil. 2005. “Structural Equations, Treatment Effects, and Econometric Policy Evaluation.” *Econometrica* 73(3): 669–738.

Imbens, Guido W., and Joshua D. Angrist. 1994. “Identification and Estimation of Local Average Treatment Effects.” *Econometrica* 62(2): 467–475.

Imbens, Guido, Donald Rubin, and Bruce Sacerdote. 2001. “Estimating the Effect of Unearned Income on Labour Earnings, Savings and Consumption: Evidence from a Survey of Lottery Players” *American Economic Review* 91(4): 778–94.

Kaboski, Joseph J., and Robert M. Townsend. 2011. “A Structural Evaluation of a Large-Scale Quasi-Experimental Microfinance Initiative.” *Econometrica* 79(5): 1357–1406.

Keane, Michael P., and Kenneth I. Wolpin. 1997.

“The Career Decisions of Young Men.” *Journal of Political Economy* 105(3): 473–522.

Koujianou-Goldberg, Pinelopi. 1995. “Product Differentiation and Oligopoly in International Markets: The Case of the U.S. Automobile Industry.” *Econometrica* 63(4): 891–951.

Lee, Donghoon, and Kenneth I. Wolpin. 2006. “Intersectoral Labor Mobility and the Growth of the Service Sector.” *Econometrica* 74(1): 1–46.

Low, Hamish. 2005. “Self-Insurance in a Life-Cycle Model of Labour Supply and Savings.” *Review of Economic Dynamics* 8(4): 945–75.

Low, Hamish, Costas Meghir, and Luigi Pistaferri. 2010. “Wage Risk and Employment Risk over the Life Cycle.” *American Economic Review* 100(4): 1432–67.

Low, Hamish, and Luigi Pistaferri. 2015. “Disability Insurance and the Dynamics of the Incentive–Insurance Trade-off.” *American Economic Review* 105(10): 2986–3029.

MaCurdy, Thomas E. 1983. “A Simple Scheme for Estimating an Intertemporal Model of Labor Supply and Consumption in the Presence of Taxes and Uncertainty.” *International Economic Review* 24(2): 265–89.

Magnac, Thierry, and David Thesmar. 2002. “Identifying Dynamic Discrete Decision Processes.” *Econometrica* 70(2): 801–816.

McFadden Daniel. 1989. “A Method of Simulated Moments for Estimation of Discrete Response Models without Numerical Integration.” *Econometrica* 57(5): 995–1026.

Meghir, Costas, and Steven Rivkin. 2011. “Econometric Methods for Research in Education.” Chap. 1 in *Handbook of the Economics of Education*, edition 1, vol. 3, edited by Erik Hanushek, Stephen Machin, and Ludger Woessmann. Elsevier.

Meghir, Costas, and Guglielmo Weber. 1996. “Intertemporal Nonseparability or Borrowing Restrictions? A Disaggregate Analysis using a U.S. Consumption Panel.” *Econometrica* 64(5): 1151–81.

Miranda, Mario J., and Paul L. Fackler. 2002. *Applied Computational Economics and Finance*. MIT Press

Orcutt, Guy H., and Alice G. Orcutt. 1968. “Incentive and Disincentive Experimentation for Income Maintenance Policy Purposes.” *American Economic Review* 58(4): 754–72.

Pakes, Ariel, and David Pollard. 1989. “Simulation and the Asymptotics of Optimization Estimators.” *Econometrica* 57(5): 1027–57.

Rosenzweig, Mark R., and Kenneth I. Wolpin. 1980. “Testing the Quantity–Quality Fertility Model: The Use of Twins as a Natural Experiment.” *Econometrica* 48(1): 227–40.

Ruge-Murcia, F. 2012. “Estimating Nonlinear

DSGE Models by the Simulated Method of Moments: With an Application to Business Cycles.” *Journal of Economic Dynamics and Control* 36(6): 914–38.

Rust, John. 1987. “Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher.” *Econometrica* 55(5): 999–1033.

Rust, John. 1992. “Do People Behave According to Bellman’s Principle of Optimality.” Working Papers in Economics E-92-10, Hoover Institution Stanford.

Schultz, T. Paul. 2004. “School Subsidies for the Poor: Evaluating the Mexican Progreso Poverty Program.” *Journal of Development Economics* 74(1): 199–250.

Su, Che-Lin, and Kenneth L. Judd. 2012.

“Constrained Optimization Approaches To Estimation of Structural Models.” *Econometrica* 80(5): 2213–30.

Todd, Petra E., and Kenneth 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* 96(5): 1384–1417.

Voena, Alessandra. 2015. “Yours, Mine, and Ours: Do Divorce Laws Affect the Intertemporal Behavior of Married Couples?” *American Economic Review* 105(8): 2295–2332.

Zeldes, Stephen P. 1989. “Consumption and Liquidity Constraints: An Empirical Investigation.” *Journal of Political Economy* 97(2): 305–46.

This article has been cited by:

1. Michael G. Findley, Kyosuke Kikuta, Michael Denly. 2021. External Validity. *Annual Review of Political Science* **24**:1, 365-393. [[Crossref](#)]
2. Chunrong Ai, Oliver Linton, Zheng Zhang. 2021. Estimation and inference for the counterfactual distribution and quantile functions in continuous treatment models. *Journal of Econometrics* **10**. . [[Crossref](#)]
3. Sadia Ilyas, Rajiv R. Srivastava, Hyunjung Kim, Subhankar Das, Vinay K. Singh. 2021. Circular bioeconomy and environmental benignness through microbial recycling of e-waste: A case study on copper and gold restoration. *Waste Management* **121**, 175-185. [[Crossref](#)]
4. Sebastian Galiani, Juan Pantano. 2021. Structural Models: Inception and Frontier. *SSRN Electronic Journal* **80**. . [[Crossref](#)]
5. Markku Maula, Wouter Stam. 2020. Enhancing Rigor in Quantitative Entrepreneurship Research. *Entrepreneurship Theory and Practice* **44**:6, 1059-1090. [[Crossref](#)]
6. Jeppe Druedahl. 2020. A Guide on Solving Non-convex Consumption-Saving Models. *Computational Economics* **31**. . [[Crossref](#)]
7. Francisco Ceballos, Miguel Robles. 2020. Demand heterogeneity for index-based insurance: The case for flexible products. *Journal of Development Economics* **146**, 102515. [[Crossref](#)]
8. Martin J Williams. 2020. External Validity and Policy Adaptation: From Impact Evaluation to Policy Design. *The World Bank Research Observer* **35**:2, 158-191. [[Crossref](#)]
9. Helena Isidro, Dhananjay (DJ) Nanda, Peter D. Wysocki. 2020. On the Relation between Financial Reporting Quality and Country Attributes: Research Challenges and Opportunities. *The Accounting Review* **95**:3, 279-314. [[Crossref](#)]
10. Ramses Armendariz. 2020. Hunger and the Demands for Cheap Dietary Staples of the Extreme Poor. *SSRN Electronic Journal* **79**. . [[Crossref](#)]
11. Arno Parolini, Wei Wu Tan, Aron Shlonsky. 2019. Decision-based models of the implementation of interventions in systems of healthcare: Implementation outcomes and intervention effectiveness in complex service environments. *PLOS ONE* **14**:10, e0223129. [[Crossref](#)]
12. Eszter Czibor, David Jimenez-Gomez, John A. List. 2019. The Dozen Things Experimental Economists Should Do (More of). *Southern Economic Journal* **86**:2, 371-432. [[Crossref](#)]
13. I P Kinash, L M Arkhypova, A S Polyanska, O G Dzoba, U Y Andrusiv, Iu I Iuras. 2019. Economic evaluation of tourism infrastructure development in Ukraine. *IOP Conference Series: Materials Science and Engineering* **477**, 012020. [[Crossref](#)]
14. Xintong Han, Pu Zhao. 2019. Pay for Content or Pay for Marketing? An Empirical Study on Content Pricing. *SSRN Electronic Journal* . [[Crossref](#)]
15. Xintong Han, Zimin Liu, Tong Wang. 2019. Multidimensional Non-Linear Pricing: Evidence from Energy Consumption with a Mixture Pricing Mechanism. *SSRN Electronic Journal* **67**. . [[Crossref](#)]
16. Olivier Bargain. 2017. Welfare analysis and redistributive policies. *The Journal of Economic Inequality* **15**:4, 393-419. [[Crossref](#)]
17. Karthik Muralidharan, Paul Niehaus. 2017. Experimentation at Scale. *Journal of Economic Perspectives* **31**:4, 103-124. [[Abstract](#)] [[View PDF article](#)] [[PDF with links](#)]
18. Arno Parolini, Wei Wu Tan, Aron Shlonsky. 2017. A Blueprint for Causal Inference in Implementation Systems. *SSRN Electronic Journal* . [[Crossref](#)]