STRUCTURAL, EXPERIMENTALIST, AND DESCRIPTIVE APPROACHES TO EMPIRICAL WORK IN REGIONAL ECONOMICS*

Thomas J. Holmes
Federal Reserve Bank of Minneapolis, National Bureau of Economic Research, Department of Economics, University of Minnesota, 4-101 Hanson Hall, 1925 Fourth Street South, Minneapolis, MN 55455. E-mail: holmes@umn.edu

ABSTRACT. The three general approaches to empirical work in economics are structural, experimentalist, and descriptive. This paper provides an overview of how empirical work in regional economics fits into these three categories. In particular, I examine a single issue in the field, the nature of agglomeration benefits and the productivity gains from agglomeration, and analyze the advantages and drawbacks of following each of these three empirical approaches. I also discuss potentially fruitful ways empirical work in regional economics might advance.

1. INTRODUCTION

The three general approaches to empirical work in economics are structural, experimentalist, and descriptive. In this paper, I provide an overview of how empirical work in regional economics fits into these three categories. I also discuss potentially fruitful ways in which the field might advance. But first, let me briefly define each category.

The *structural approach* in empirical economics starts with a fully specified economic model, well grounded in theory. The goal of this approach is the estimation of the underlying "deep model parameters" of preferences and technology. Once the model parameters are obtained, we have an artificial economy in hand that we can put to work, simulating the impact of various policy alternatives. Importantly, we can even use the estimated model to study the impacts of policies that have never yet been implemented. This approach is the dominant paradigm for empirical work in the field of industrial organization (IO) (see Ackerberg et al., 2007, for a primer). The approach has also made

*The research here was funded by National Science Foundation Grant SES 0551062. I have benefited greatly from the comments of the special editor Gilles Duranton, the discussant Patrick Bayer, and the participants at the *Journal of Regional Science*'s 50th Anniversary Conference held at the Federal Reserve Bank of New York. The views expressed herein are those of the author and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System.
Received: April 2009; revised: July 2009; accepted: July 2009.

 DOI: 10.1111/j.1467-9787.2009.00637.x

headway in labor economics (Keane and Wolpin, 1997; Eckstein and Wolpin, 1999) and in public economics (Epple and Sieg, 1999).

The structural approach can be contrasted with the *experimentalist approach* that originated in the labor economics literature.[1] In this approach, the identification of the causal impact of a policy treatment is the goal. But rather than look through the lens of economic theory, the approach relies on finding "natural experiments" or clever instruments to tease out the impacts of policies that have already taken place. Finding these experiments enables the researcher to treat behaviors as exogenous that would otherwise have to be treated as endogenous. For example, Angrist and Krueger (1991) tease out a "return to education" from the fact that when the minimum age for dropping out of high school is 17, a teenager who turns 17 in November rather than August has a "treatment" of three more months of mandatory education. The experimentalist approach often makes use of a regression discontinuity (RD) design (Hahn, Todd, and van der Klaauw, 2001). Consider a university that awards scholarships to students who score 95 or above on a test. A study of the effects of the scholarship that compares students scoring 95.1 on the test (and being awarded the scholarship) with students scoring 94.9 (those essentially as smart as the students in the other group but just missing the scholarship) makes use of an RD design. In the literature, the first group is the "treated group" and the second the "control" in an "experiment."[2]

The structural and experimentalist approaches are similar in that both seek to identify quantitative impacts of policy changes. In this way, they differ from the *descriptive approach* that can make no such claim to identification. Papers following the descriptive approach usually begin with some discussion of economic theory and empirical implications that can be derived from the theory, the latter typically in the form of correlations or monotonicities we expect to observe when one variable is plotted against another. The signs and statistical significance of empirical relationships are often the main focus of the empirical analysis (as opposed to the magnitude), as the goal is usually to discriminate across alternative theories that differ in the signs of comparative statics relationships. Regression parameter estimates obtained in this style of work are typically referred to as *reduced-form* estimates, as opposed to underlying structural parameter estimates.

Traditionally, much applied work in urban and regional economics has followed the descriptive approach. In recent years, experimentalist ideas have begun to be adopted by the field. Analysis of the impact of government policies on business location is a natural application of RD design because government policies change discontinuously at political unit boundaries. See Holmes (1998) and Duranton, Gobillon, and Overman (2007) for studies along this line. Experimentalist ideas have been used to study the impact of agglomeration. Clever

[1]See Angrist and Krueger (1999) for a survey of the experimentalist approach. See Keane (2009) for a critical view.

[2]See van der Klaauw (2002) for such a study of the impact of financial aid.

ideas that have been considered include the natural experiment of variations in the location of bedrock in Manhattan for supporting tall buildings (Rosenthal and Strange, 2008) and the bombing of Hiroshima (Davis and Weinstein, 2002). In contrast, the structural approach has not made much headway in urban and regional economics except for the corner of the field that overlaps with local public finance, such as Epple and Sieg (1999).

This paper argues that all three approaches have a role to play. Natural experiments are great if you can find them, but they are generally not as easy to come by for questions in regional economics as compared to labor economics. This suggests a potentially large role for economic theory as a substitute for experiments in disentangling what is going on in the data. That is, the structural approach has much to offer to the field.

I make these points by examining a single issue in the field: the nature of agglomeration benefits and the productivity gains from agglomeration. This is a vast literature, and I make no attempt to survey it (see Rosenthal and Strange, 2004, instead for a comprehensive survey). Rather, I select a few papers that are useful for making the comparisons across empirical approaches that I want to highlight.

2. DESCRIPTIVE WORK ON AGGLOMERATION

Agglomeration of economic activity—whether it be a large population of individuals forming a city or a large group of plants in the same industry forming an industry district—is a core subject of interest in urban and regional economics. Economic units agglomerate to a remarkable extent. How important a role do natural location characteristics play in agglomeration (i.e., Ricardian productivity advantages), and how important a role do scale economies play? To the extent that agglomeration is driven by scale economies, what is the nature of the scale economies? Are the scale economies external or internal? These are big questions in the field.

Much empirical work on agglomeration is descriptive. The results take the form of stylized facts in cross tabulations and descriptive regressions, or even in maps like the one in Burchfield et al. (2006), which quantifies agglomeration and land cover based on satellite pictures from space.

As an illustrative example of this kind of work, let me briefly discuss one of my own papers (Holmes, 1999). A leading theory for why a particular industry might "localize" in one place is that agglomeration facilitates the emergence of a wide diversity of locally provided intermediate inputs to the industry in the area. With access to a wide variety of local suppliers, in theory we expect plants in such areas to more readily outsource tasks to local suppliers rather than vertically integrate and do the tasks themselves, as compared to plants in isolated areas. Holmes (1999) examines this hypothesis using data from the United States manufacturing sector. Looking within narrowly defined industries (the four-digit SIC code level), the paper constructs a measure of vertical integration and compares the average value of this measure among plants in high industry concentration areas (i.e., plants for which there is a large amount of employment in the same four-digit industry in nearby plants) with plants in low industry concentration areas. The main finding is that there is a clear pattern: the vertical integration measure within an industry is indeed lower in high concentration areas. Li and Lu (2009) corroborate this finding for China, and Ono (2003) reports a related finding.

Establishing a basic fact like this in a broad data set is sensible. A qualitative finding such as this can guide future theoretical efforts. Furthermore, before jumping into any kind of complicated structural estimation, we should examine the patterns in the raw data as a first step. Nevertheless, to make further progress in the field, going beyond descriptive work is useful for two reasons.

First, descriptive work like Holmes (1999) does not deliver estimates of model parameters. We do not get an economic model that we can put on the computer to analyze the impact of policies.

Second, the paper does not directly confront a key selection issue. If there are substantial benefits for a plant to be located in an industry concentration center, why are plants ever located in an isolated area? Plants in isolated areas may differ in systematic ways from plants in industry concentration centers. If so, when we compare them, it is not an everything-else-held-fixed comparison. It is an apples-to-oranges comparison rather than apples-to-apples.

Holmes (1999) raises these issues and suggests that there may potentially be factors on both the supply and demand sides driving plants to open in remote locations. On the supply side, entrepreneurs with particular skills in a particular industry might end up in locations outside of industry concentration centers for various reasons (e.g., to follow a spouse's career). On the demand side, a component of demand in a given area might for some reason be accommodated only by a local provider, and so a producer might locate there even if the location is far from an industry center. It may be possible to tell a story for why these supply or demand factors might be exogenous in econometric terms so that the plants in concentrated and isolated areas would be similar, on average, in their response to opportunities to vertically disintegrate, that is, it is possible that this is a *ceteris paribus* comparison.

But more likely than not, the plants in isolated locations differ in other respects compared to plants in concentrated areas. In recent work, Holmes and Stevens (2009) argue that even within the most narrowly defined Census industrial classifications, there tends to be great heterogeneity in what plants do that is systematically related to whether or not the plants are located in industry concentration centers. In particular, plants in industry centers tend to specialize in mass production goods with relatively low transportation costs, whereas plants in the same industry outside of industry centers tend to specialize in custom goods meant for local consumers. Consider the wood furniture industry, for example. In the United States, the industry is centered in North Carolina where factories make stock furniture pieces through mass production techniques. The same Census industry includes craft furniture shops scattered throughout the country that make custom furniture to order for local consumers. A comparison of mass production factories in North Carolina with craft shops scattered around the United States is not an apples-to-apples comparison. In this way, such comparisons give potentially biased estimates of how an increase in localization would impact the vertical integration decisions of a particular plant.

A related issue raised in Rosenthal and Strange (2004) is that plants within a given industry might be heterogeneous in the value they place on access to local suppliers. So plants placing a high premium on access to local suppliers might disproportionately sort into locations with high industry concentration and lots of suppliers. On account of this sorting effect, we might expect the comparison of isolated plants with high industry concentration plants to bias upward the impact of localization on an individual plant's decision to vertically disintegrate.

The issues I am raising here about my own paper can be raised about all descriptive work on industry localization. If there are advantages to localization, what factors lead plants to open in areas without the benefits of localization? And what can be learned by comparing such plants with those choosing to locate in high concentration areas?

3. THE EXPERIMENTALIST APPROACH

The experimentalist approach attempts to directly confront the issues of selection and endogeneity. I illustrate this approach by discussing how it has been used to estimate the impact of agglomeration on wages.

Consider an environment where workers are paid their marginal product, and for simplicity ignore capital. Suppose the log wage of worker j at location ℓ at time t obeys the following relationship:

$$\ln w_{j\ell t} = \beta_t + \gamma \ln N_{\ell t} + \alpha_\ell + \theta_j + \varepsilon_{jt}. \tag{1}$$

The first term β_t is a fixed time effect. The parameter γ is the crucial parameter governing the agglomeration benefit. Defining $N_{\ell t}$ to be the population at location ℓ at time t, γ is the elasticity of wages (equivalent to productivity here) with respect to agglomeration. The third term α_ℓ, a fixed location effect, captures the contribution of a location's features to a worker's productivity. The fourth term θ_j is a fixed effect associated with worker j, that is, ability. Finally, assume the last term ε_{jt} is a random, i.i.d. shock to worker j's wages at time t.

For simplicity, let us begin by assuming that θ_j is directly observable and bring this term over to the left-hand side. In addition, suppose we have just one period of data, $t = 0$. Then we can rewrite the above as

$$\ln w_{j\ell 0} - \theta_j = \beta_0 + \gamma \ln N_{\ell 0} + \alpha_\ell + \varepsilon_{j0}. \tag{2}$$

Suppose then that we regress worker log wage net of ability, $\ln w_{j\ell t} - \theta_j$, on a constant and the log population ($\ln N_{\ell t}$) where the individual lives. Then $\alpha_\ell + \varepsilon_{j0}$ is an error term. This regression exercise is a good example of descriptive empirical work. From this exercise, we learn the cross-sectional correlation between wages (net of ability) and agglomeration at time $t = 0$. But we do not expect that this ordinary least squares (OLS) exercise identifies the structural coefficient γ. This follows because we expect that the level of agglomeration $N_{\ell 0}$ is endogenous and positively correlated with the location productivity term α_ℓ. Everything else the same, we expect more productive locations to attract more workers. Hence, the regressor $\ln N_{\ell 0}$ is positively correlated with the error term $\alpha_\ell + \varepsilon_{j0}$ in Equation (2), so OLS yields an inconsistent estimate of γ.

To obtain a consistent estimate of γ, Ciccone and Hall (1996) propose the use of instrumental variables. (Their data are more aggregated than individual worker data, but the logic is the same as for this case.) In particular, as an instrument for $\ln N_{\ell 0}$, they propose the use of long lagged values of population (over 100 years!). They argue first that population today $N_{\ell,0}$ is correlated with population from 100 years ago $N_{\ell,-100}$ because there is inertia for where people live. They argue second that today's location productivity characteristics, captured by α_ℓ, are unlikely to be correlated with population from 100 years ago $N_{\ell,-100}$ because the economy has undergone such dramatic change over the course of a century. To the extent that both claims are true, long lagged population is a valid instrument. The allocation of population 100 years ago, which is claimed to be random from the perspective of today's productivity considerations, is the "experiment" that identifies γ.

Combes, Duranton, and Gobillon (2008) is a recent paper to take this approach. In addition to addressing the problem of endogeneity of population $N_{\ell t}$ through instrumental variables (IV), the paper also confronts the issue that worker ability θ_j has an unobservable component. Because we expect workers to sort themselves across locations, this means there is a problem of selection. Suppose in the original regression (1) that θ_j is unobservable and that high-ability people tend to sort into high population areas. Then we see two reasons why an OLS regression of (1) does not identify γ: (1) the level of population $N_{\ell t}$ is correlated with location productivity α_ℓ, and (2) worker ability θ_j is correlated with population. To resolve the selection problem caused by worker sorting, Combes et al. (2008) follow Glaeser and Mare (2001) by using a panel approach. Observing workers who switch locations over time, they can difference out the worker fixed effect θ_j to estimate the remaining parameters, and then recover the θ_j in a final step. Combes et al. (2008) tackle the issue with an extremely rich and detailed data set with which they can track a huge sample of workers in France over a long period of time. Their preferred estimate of γ is approximately 0.03. This is less than that of some other papers in the literature, and they attribute the difference to the fact that they take sorting into account. Combes et al. (2010) push this agenda further by incorporating additional instruments for population in the form of soil quality measures.

One can criticize the use of long population lags as an instrument for current population. In particular, one can try to make the case that some of the same geographic features (such as a harbor) that made a place productive

100 years ago may still play a role today. Nevertheless, as things go with papers in the IV literature, the case for the instrument here is more compelling than usual.

A difficulty that comes up with this approach, however, is that it does not "scale up" so easily to look at the impact of more narrowly defined agglomerated activities than population as a whole. To make this point, generalize Equation (1) to allow for an agglomeration benefit that depends upon the industry i an individual worker is in. Specifically, suppose wage equals

$$\ln w_{j\ell ti} = \beta_t + \gamma \ln N_{\ell t} + \lambda_i \ln M_{j\ell ti} + \alpha_\ell + \omega_{\ell i} + \theta_j + \varepsilon_{jt}, \tag{3}$$

where $M_{j\ell ti}$ is employment at location ℓ in industry i at time t where person j works and $\omega_{\ell i}$ is the productivity of industry i at location ℓ. In the literature, the coefficient on total population γ is the *urbanization* coefficient, whereas the coefficient λ_i on the own industry level is the *localization* coefficient for industry i.

By the same logic that we might have endogeneity of population $N_{\ell t}$ (responding to differences in α_ℓ across locations), we expect there to be endogeneity of industry employment $M_{\ell ti}$ (responding to differences in the location/industry term $\omega_{\ell i}$). In principle, we can try to use long lags in industry employment to instrument for current industry employment, by analogy of the instrument for total population. An immediate difficulty is that many industries we would be interested in (e.g., semiconductors) did not exist 100 years ago. Rather than use very long lags, we could consider using short lags. The danger is that with short lags, we run the risk that the same location factors driving industry location in the lagged period are still at work in the current period, invalidating the short lag as an instrument. The presence of Stanford University in Silicon Valley in California was a key reason the semiconductor industry began in this area in the 1950s, and the impact of Stanford on this industry continues to this day. Combes et al. (2008) allow for a localization term as in Equation (3). They recognize that they do not have instruments for individual industry localization and so cannot correct for endogeneity of $M_{\ell ti}$. Instead, they focus on what they can fix, selection on θ_j and endogeneity of $N_{\ell t}$. Surely in some example industries, the use of a particular level of lagged employment as an instrument makes sense. But these need to be considered on a case-by-case basis rather than used wholesale across a large set of industries. With specific knowledge of an industry, a researcher can utilize as an instrument a particular factor that was once a key driver of location in an industry but is now irrelevant because of technological change. For example, Arzaghi and Henderson (2008) examine location patterns in the advertising industry and argue that the proximity to client corporate headquarters was once such a key location but is less important now.

The logic underlying the above resolution of the endogeneity of $N_{\ell t}$ is the cross-sectional variation of the instrument. One hundred years ago, dramatic differences could be found in population levels *across* locations, and dramatic differences persist to this day. An alternative strategy is to estimate the impact of localization by relying on variation *within* a location over time, which enables the researcher to hold fixed permanent unobservable location characteristics (such as proximity to Stanford). Henderson (2003) is a leading recent example pursuing this strategy. A limitation of the *within* approach is that agglomeration of an industry at a location tends to change relatively slowly over time, so the within variation is relatively small compared to the dramatic cross-sectional variation. If local industry scale increases by 10 percent, and if the localization coefficient is $\lambda = 0.03$, the predicted increase in output is 0.3 percent. This may be hard to tease out, given substantial measurement issues and the difficulties of accounting for other factors that change. Moreover, the temporal connection is subtle. If a factory hires 10 percent more workers and uses 10 percent more materials in a day, we might expect to see something like 10 percent more output coming out of the factory that very day. But the spillover impact on the output of other plants is more likely to take place with a lag. Given such potential lags, the spillovers are likely easier to pick up in a cross-sectional comparison of long-run situations rather than through a within comparison of short-run fluctuations. Finally, researchers need to confront the issue that changes over time in $M_{\ell i}$ are likely to be closely connected to shocks over time to a location's suitability $\omega_{\ell i}$ for an industry. Henderson (2003) addresses this point by using dynamic panel techniques in which variables in different form can be instrumented for by sufficiently long lags of variables in level form. Intuitively, if there is a form of "regression to the mean" in the underlying data-generating process, then high levels of lagged employment may be correlated with changes (declines) in current employment levels, but not correlated with current productivity shocks. However, Henderson (2003) does not get much mileage out of this strategy because these instruments turn out to be weak in the application. The key step of the experimentalist approach is always how to come up with good instruments.

It can be particularly difficult to come up with instruments for localization compared to other applications. Consider an application where a researcher is interested in whether or not high school students learn to be a criminals when sitting next to classmates who are already criminals.[3] Compared to a typical localization application, there is vast asymmetry in terms of the possible frequency count of data one might hope to get for such research. There are millions of high school students, and the process repeats over and over every year with new high school students. Contrast this with the topic of localization benefits of the semiconductor industry, where just a few outlier locations have attracted the vast bulk of the industry. Moreover, the movement of this industry is not something that happens every year, but is a process that involves decades. Now let us say a researcher has a clever strategy for an instrument, but it only works in a fraction 0.0001 of observations. Given the millions of times the event of a student sitting next to another student is repeated, a researcher may have

[3]On a more positive note, suppose instead individuals can learn about work opportunities from people with whom they interact (see Bayer, Ross, and Topa, 2008).

reason to hope that he or she can find a large subset of observations for which the instrument applies. But let us say for a localization application there is a location A with a large amount of the industry and a location B with a little, and no other location has any. If the chance of finding a good instrument is 0.0001 per observation, the researcher is in trouble.

Researchers in regional economics interested in using the experimental approach might find it useful to flip the approach around somewhat to make things more like a labor market application. In particular, rather than treat the source of spillovers as occurring at the level of a location as a whole, it may be better to track the origins of the spillovers plant by plant. When we work at the industry level, supposing there are just two locations A and B, we may not have an instrument for why the industry at the aggregate level has tended to be at A rather than B. But suppose instead we can look at the plant level and we have particular information about new plant openings. Suppose we can procure an instrument for the location decisions of individual plants. To the extent we set this up as an experiment so that we can interpret A as treated with a plant and B as a control getting no plant, we can identify the impact on the productivity of the neighboring plants at location A. Greenstone, Hornbeck, and Moretti (2008) is an example of this approach. They obtain site selection information about new plants, including in particular information about the first and second choice locations of each plant. They set up a regression discontinuity design as an attempt to treat the second-choice locations (that did not get the plant) as a control group. The locations just missing getting the plant are the analog of the students discussed in the introduction who just missed getting the scholarship.

Analogous to the way we might want to disaggregate the *sources* of spillovers, we might also want to consider disaggregating the *recipients* of the spillovers. There is only one Silicon Valley, and if all recipients of spillovers are the same, there may be little we can say. But if there is heterogeneity in the ability to enjoy spillovers, and if we can somehow get a data handle on this characteristic, we can make progress by looking across firms within Silicon Valley at a point in time that vary in this characteristic (as opposed to comparing Silicon Valley with another place or Silicon Valley with itself at a different point in time). And by progress, I mean that with this alterative source of variation, we can potentially better address the same questions considered before, as well as address new questions. Ellison, Glaeser, and Kerr (2007) provide an analysis in this spirit that looks at coagglomerated industries—different industries that tend to be grouped together in space. As these industries have different measurable characteristics (i.e., different input–output requirements) and value different things, this heterogeneity can be exploited to shed light on what makes locations drawing coagglomerated industries so attractive. Holmes (2005), discussed later, is another example of a paper using measured firm heterogeneity to recipients of agglomeration benefits (in this case, the measured heterogeneity is firm size) to disentangle the sources of benefits from agglomeration.

4. THE STRUCTURAL APPROACH

A researcher taking the structural approach begins with a fully specified economic model that specifies the forces driving the variation in the data and the error terms. A cost of the structural approach is that the researcher must be explicit up front about the underlying model and underlying economics. (But see Keane, 2009, for an argument that experimentalists actually take stands up front as well, only implicitly.) In return for being explicit up front about the model, a researcher taking the structural approach potentially gets two benefits. First, through use of the theory, the researcher can obtain restrictions to help pin down parameters. Second, by estimating the structural parameters of a fully specified model, the researcher can simulate the impact of alternatives that have never actually occurred in the data.

On the Use of Structure to Pin Down Parameters

In scientific inquiry where controlled experiments are feasible, it may be possible to get quite far without any theory. A medical researcher giving a treatment to one set of subjects and a placebo to a control group can estimate the physiological effects of the treatment without necessarily developing a theory of how it works. The actual mechanics of what is happening in the body can be a "black box."

Controlled experiments do take place in economics. There is now a whole field devoted to running lab experiments typically related to topics in game theory or questions about how prices get determined in market situations. In labor economics, social experiments are actually run (e.g., randomized trials of work subsidies or educational opportunities). But in regional economics, there is little possibility for running controlled experiments. Good luck to the researcher trying to get a National Science Foundation grant to run a controlled experiment to set up a new city! So researchers obviously have to make the best of things working with data not under their control. One strategy for making the best of things is to bring theory into the analysis.

Theory plays little role in the work on wages and agglomeration discussed in the previous section. To be precise, there is actually a bit of theory in the beginning of these papers, where a wage equal to marginal product condition is derived from profit maximization. But theory does not play a role further in. The papers ask how a worker's wage would change if, experimentally, the worker were moved from a smaller city into a larger city. (Or if, experimentally, the city he or she worked in was made bigger.) The answer is obtained by trying somehow to find the direct analogs of these experiments in the data.

Several recent leading papers on agglomeration do feature economic theory prominently in the analysis, using theory to set up a choice problem for a firm location decision (Rosenthal and Strange, 2003; Arzaghi and Henderson, 2008). These papers use restrictions implied by the choice model to motivate a reduced-form entry equation, and they estimate the parameters of the reduced form. Among other results, their estimates show how entry varies with distance to agglomerations (with distance delineated in relatively fine increments). While these papers very much appeal to theory, they do not go all the way in the full structural agenda to produce estimates of a complete economic model with all the underlying technology parameters specified. Interestingly, there are developments in the structural literature in labor and IO where reduced-form estimations of policy functions (such as an entry equation) play crucial roles as a first step of the analysis. In Hotz and Miller (1993) and Bajari, Benkard, and Levin (2007), reduced-form behavioral relationships are estimated in a first stage, and in a second stage the structural parameters are recovered that would generate these behavioral relationships.[4] (Breaking the problem down this way solves certain complexity issues that arise in dynamic environments, particularly with multiagent decision making.) In this way, we can think of estimates of reduced-form behavioral relationships as potentially complementing future structural work, being a first stage, rather than a substitute for structural work.

I now review two of my own studies on agglomeration that take the structural approach and in each case illustrate the role that theory plays in pinning down parameters. In both cases, the basic theory that is exploited is revealed preference. By observing what agents are choosing given their choice sets, and assuming rational maximizing behavior, drawing inferences about the underlying parameters is possible. This is, of course, very different from a controlled experiment where a subject mechanically gets a treatment and is not making any choices.

Holmes (2005) examines the concentration of manufacturing sales offices in large cities. Sales offices are the home bases of company representatives who call on customers to mediate sales. This activity is a good example of the kind of white-collar, information-intensive work that tends to be the provenance of large cities. We might expect large cities to attract this line of work for two main reasons. First, big cities obviously have high market potential compared to small cities. A sales representative located in a very large city has many nearby local clients to keep busy with, compared to a sales representative working out of a small city. Second, workers in big cities may be potentially more productive, both because of knowledge spillovers and because high-skill workers may sort into big cities. Holmes (2005) uses confidential Census micro data to estimate a structural model of sales office location choice that disengages the relative importance of the market potential and productivity factors. Each manufacturing firm chooses a network of locations to place sales offices, taking these two considerations into account and also taking into account the fixed costs of setting up sales offices. The main idea underlying the parameter

[4]For example, in a first stage, a model of the probability of new entry into a market might be estimated that depends upon the market size, the number of incumbents, and other market-level variables that are observable to market partipants and an econometrician analyzing the data. In a second stage, structural parameters (e.g., the distribution of entry costs) are estimated that would rationalize these first-stage behavioral relationships.

identification is that manufacturing firms of different sizes will respond differently to the market potential and productivity factors. (This is an example of the strategy mentioned in the previous section to exploit measurable differences among the recipients of agglomeration benefits to disentangle the nature of these benefits.) The differences arise because large firms have the scale economies to open up a large network of sales offices covering most of the national market, so large firms can expect to get large market access whatever they do. Small firms will open just a few offices and will get relatively low market access whatever they do. The paper shows how these firm size considerations "tilt" a firm's indifference curves over possible network configurations a certain way. By tracing out how these indifference curves shift with firm size, it is possible to back out the structural parameters underlying these indifference curves. The main finding is that the market potential factor is large compared to the productivity factor in accounting for location decisions in this industry.

Theory plays a crucial role in this analysis. It delivers a rich set of moment conditions that identify the underlying parameters. There does not appear to be any way to come up with a single-equation reduced-form behavioral equation summarizing all the content of the theory. The theory has some *qualitative* implications, and the paper undertakes descriptive analysis of the raw data to verify that these qualitative implications indeed hold. But the quantitative claim of the paper about the relative importance of the market potential and productivity factors rests entirely on the structural analysis. It is possible to imagine a study on this topic taking the experimentalist approach. But I believe it would require more data than I use here, and in particular would have to include dollar valuations on sales representative traveling time costs, which would be difficult to measure (especially since the costs more generally include the degradation in service quality when a major account is serviced by an out-of-town sales office). Finally, I note that the analysis treats firm size as an instrument (under the argument that some firms are huge and others small for reasons mostly unrelated to the particular way the firms have set up their sales office structures). So the difference between the structural approach and the experimentalist approach is not the subtraction of instruments, but rather the addition of theory.

Next consider the Holmes (2008) paper on the Wal-Mart store location problem. The question addressed is: How big are the cost savings enjoyed by Wal-Mart when it "agglomerates" its stores and packs them closely together? By maintaining a dense network of stores, Wal-Mart can run an efficient logistics system with distribution centers placed very close to stores. This strategy keeps trucking costs low and also allows Wal-Mart to quickly replenish its store inventories in response to demand shocks. Wal-Mart famously was able to quickly restock its shelves with American flags after 9/11. One can readily imagine an experimentalist approach to estimating the benefit of store agglomeration that regresses store costs on store density, perhaps using population density as an instrument. The structural approach used in Holmes (2008) has three advantages over such an experimentalist approach. First, Wal-Mart is a highly secretive company and is unlikely to ever provide store-level cost data to researchers. The advantage of the structural approach is that it uses the information revealed by the choice behavior of Wal-Mart as a substitute for the unobtainable data. Second, even if accounting data were available directly from Wal-Mart, it might be hard to quantify the benefits to Wal-Mart of its maintaining its quick replenishment strategy. (This is analogous to the difficulty in quantifying the cost of serving a major account by an out-of-town sales office.) But again, through the revealed preference of Wal-Mart's choices, it is possible to infer what the benefits must be worth to Wal-Mart. Third, population density turns out to be an invalid instrument because it is correlated with other aspects of costs. The structural approach makes it possible to jointly estimate the parameters governing how store density *reduces* costs and population density *increases* costs.

These two examples illustrate the role of single-agent decision theory—revealed preference in choice behavior—in pinning down model parameters. Additional restrictions can be obtained from the equilibrium of agent interactions. The structural work in local public finance such as Epple and Sieg (1999) very much exploits the information content in the market equilibrium. Combes et al. (2009) is a recent paper that uses the equilibrium restrictions of market equilibrium to analyze agglomeration economies. Plants in big cities tend to have higher measured productivity than plants in small cities. Is this because of an agglomeration benefit? Or is this on account of selection? (The assumption is that the selection process to weed out unproductive firms is tougher in larger markets, as highlighted in recent theoretical work.) The paper writes down an explicit model of agglomeration and selection that nests both of these explanations. Combes et al. (2009) show with the model that by comparing the distribution of productivities across cities of different sizes (arising in market equilibrium), it is possible to identify relative values of the key structural parameters. That is, the level of the agglomeration parameter γ from the earlier section is not identified, but the ratio of this parameter to the parameter governing selection is identified. The main finding is that plants have higher productivity in bigger cities primarily because of agglomeration benefits rather than the selection effect.

To follow up the earlier point that the structural approach is not about the subtraction of instruments, but rather the addition of theory, the paper by Au and Henderson (2006) is worth noting. This paper studies the efficiency of the size distribution of cities in China. It treats the size distribution like an instrument (making the case that it is exogenous). It adds an economic model of a city with which it is possible to make welfare calculations. It can therefore conduct that kind of policy analysis discussed in the next subsection.

Finally, I note the paper by Gould (2007). The structural labor literature (e.g., Keane and Wolpin, 1997; Eckstein and Wolpin, 1999) is concerned with (1) what choices individuals make and (2) how wages vary depending on these choices. Typically, the choice considered in the literature is years of education or something similar. But there is a clear analogy to think about the choice as being where to live and to consider how the choice of where to live impacts a worker's wage. Gould uses a structural labor approach to do exactly that.

On the Use of Structure to Run Policy Experiments

It is well understood in the profession that reduced-form relationships are limited in their use for policy analysis. As well articulated in Lucas (1976), reduced-form parameters are functions of underlying structural parameters and policy choices. If policy changes, in general the reduced form changes as well. In contrast, if a researcher has the parameters of a complete structural model, the researcher can put the model on the computer and simulate the impact of policy changes.

The experimentalist approach aims to identify underlying structural parameters, like the agglomeration coefficient γ in Equation (1). But in the process of estimating γ, the experimentalist approach described earlier does not try to explicitly model why people sort and live where they do. Rather, its ambition is only to "correct" for sorting or simultaneity more generally. This means that in terms of questions about policy impacts that we might consider, the experimentalist approach gives incomplete answers. If the instruments are doing their job, we can determine the treatment effect on a worker's wage of moving the worker from a small town to a large town. But we can say nothing about the costs of intervening and creating such moves.

A structural analysis is different in that it typically seeks the parameters of the whole process governing where people choose to live in addition to technology parameters governing agglomeration benefits. With the parameters of a complete model, we can put the model to work to get a complete answer to the cost and benefits of policy interventions. Moreover, if we have indeed identified the deep structural parameters that are invariant to policy changes, we can use the estimates to simulate the impact of policies that have never yet been implemented.

The application of estimated structural models to policy analysis is a prime driving force for adoption of the structural paradigm in the field of IO. For example, a simulation analysis of what happens if two firms are allowed to merge is a policy analysis that is credible only if the researcher has estimated a structural model. Analogously, the structural work in labor economics is very policy oriented. So far in regional economics (outside of the local public finance work discussed in Sieg, 2009), there is not much work exploiting structural techniques to do policy analysis. Going forward, adopting the structural approach for policy analysis is a promising direction for regional economics.

5. GOING FORWARD

I conclude with some observations and assessments about the current state of the field and where it might advance.

The experimentalist movement has been a useful development in the way that it has forced researchers to think about causality. Sometimes earlier work in the descriptive tradition would get sloppy and treat correlations in raw data as though they were causal relationships. As an intellectual movement, the experimentalist approach has been successful in that researchers do not easily get away with careless work today. In fact, papers that run simple regressions and do not address issues of causality typically face uphill battles in the publication process. So it is unnecessary for me to put in a plug for the experimentalist approach; a large installed base of advocates already does this. Instead, I take a moment to put in a plug for continued descriptive work. As noted earlier, coming up with sensible instruments can be very hard. Yet it is not uncommon for a researcher to claim that a particular endogeneity problem is solved by the use of instruments without making a strong case for both the compelling economic logic of the instruments as well as their econometric validity. Descriptive regressions of raw data and cross-tabulations establishing "stylized facts" are likely to be of more value than estimates run through an econometric meat grinder with instruments that do not make sense. There is substantial value in having the raw patterns in the data being included in our deliberations in what is happening in phenomena under investigation. It provides discipline: if patterns look very different after we run data through complicated procedures, we should understand why this is so. In particular, researchers finding a divergence between OLS and IV results should be able to account for the differences (ideally with out-of-sample evidence).

Regarding structural work, looking at where things are going in other fields to see if there are any lessons for regional economics is useful.

Regional economics is a close cousin of IO, and factors that make the structural approach successful in IO should in many cases translate into success in regional economics. In IO, the structural agenda has been put to work there mainly in two places: demand estimation and models of firm entry and exit. One core focus of regional economics is where individuals choose to live. This choice is, of course, an element of consumer demand. Epple and Sieg (1999) provide structural demand estimation that takes into account that when a household chooses to buy a house, it also chooses a school district. Bayer, Ferreira, and McMillan (2007) is another example. Another core focus of regional economics is where firms choose to enter and locate themselves. The IO literature has had large investment structural models of entry and exit, and some of the techniques may prove useful for questions in regional economics. Suzuki (2008) is a promising example that shows how zoning regulations impact the market for hotels.

One caveat should be raised regarding the comparison between IO and regional economics. It is an issue that generated much discussion at the conference where the papers in this journal issue were presented. In IO, for some applications, the model is "self-evident," at least in a relative sense. The best example is the analysis of auctions, which have specific rules. Often it is possible to write down an extensive form model of a game that the agents in an auction are playing without making too much of a leap. It is no surprise that much of the early structural work in IO focused on auctions. In contrast, the argument can be made that empirical questions in regional economics often require a bigger leap. For example, a structural model of agglomeration may contain spillovers, labor market matching, input/output linkages, and other things. It is likely that a structural model of this issue would have to abstract from potentially important aspects of the economic environment to a greater degree than the way an auction could be modeled. Hence, there is much to be said for modeling approaches like that of Combes et al. (2009), which were flexible in the way agglomeration effects operate in the model and that nest various possibilities.

Generally in applied work in economics, there is high demand for models that can be put on the computer and that (1) are fully specified rigorous economic models, (2) are rich and flexible enough to capture the key elements of the particular application in question, and (3) can be successfully fit to data in such a way that experiments one might run in the model approximate what would happen if these policies were implemented in the actual economy. The computable general equilibrium models in the trade literature and regional economics literature (Shoven and Whalley, 1984) aim to provide quantitative models to predict outcomes of trade liberalization and other shocks. In an analytically more rigorous fashion, the dynamic stochastic general equilibrium models in the macroeconomic literature (Kydland and Prescott, 1982) aim to provide quantitative models of the aggregate economy. The particular econometric criteria used for fitting these models to data is different from the way in which structural models in applied microeconomics are estimated today. But in the grand scheme of things, these earlier literatures are part of the broad structural agenda in applied work. These literatures have had a big impact. I expect that if we can create a new generation of structural models in regional economics, they will similarly have big impacts.

The discussion earlier has focused on agglomeration benefits, but progress may be made in many other directions in regional economics as well. One direction is to develop new approaches to regional input–output modeling. Current practice is based on an approach that is many decades old and does not have at its foundations a fully specified economic model. An updated approach is likely to achieve significant attention. Another direction would be to develop quantitative models of the organization of cities. An empirical tractable version of the Lucas and Rossi-Hansberg (2002) model is a direction to think about here. Finally, regional economics should pay particularly close attention to developments with its very close cousin, the field of international trade. There, new models of industry structure have been developed, such as Eaton and Kortum (2002) and Melitz and Ottaviano (2008), that are flexible and have potentially wide applicability for quantitative modeling, compared with earlier structures built on Dixit-Stiglitz foundations.

REFERENCES

Ackerberg, D., C.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*, Vol. 6A. Amsterdam: Elsevier.

Angrist, J.D. and A.B. Krueger. 1991. "Does Compulsory School Attendance Affect Schooling and Earnings?" *Quarterly Journal of Economics*, 106, 979–1014.

———. 1999. "Empirical Strategies in Labor Economics," in O. Ashenfelter and D. Card (eds.), *Handbook of Labor Economics*, Vol. 3A. Amsterdam: North-Holland, pp. 1277–1366.

Arzaghi, M. and J.V. Henderson. 2008. "Networking off Madison Avenue," *Review of Economic Studies*, 75, 1011–1038.

Au, C.-C. and J.V. Henderson. 2006. "Are Chinese Cities Too Small?" *Review of Economic Studies*, 73, 549–576.

Bajari, P., C.L. Benkard, and J. Levin. 2007. "Estimating Dynamic Models of Imperfect Competition," *Econometrica*, 75, 1331–1370.

Bayer, P., F. Ferreira, and R. McMillan. 2007. "A Unified Framework for Measuring Preferences for Schools and Neighborhoods," *Journal of Political Economy*, 115, 588–638.

Bayer, P., S.L. Ross, and G. Topa. 2008. "Place of Work and Place of Residence: Informal Hiring Networks and Labor Market Outcomes," *Journal of Political Economy*, 116, 1150–1196.

Burchfield, M., H.G. Overman, D. Puga, and M.A. Turner. 2006. "Causes of Sprawl: A Portrait from Space," *Quarterly Journal of Economics*, 121, 587–633.

Ciccone, A. and R.E. Hall. 1996. "Productivity and the Density of Economic Activity," *American Economic Review*, 86, 54–70.

Combes, P.-P., G. Duranton, and L. Gobillon. 2008. "Spatial Wage Disparities: Sorting Matters!" *Journal of Urban Economics*, 63, 723–742.

Combes, P.-P., G. Duranton, L. Gobillon, and S. Roux. 2010. "Estimating Agglomeration Economies with History, Geology, and Worker Effects," in E.L. Glaeser (ed.), *The Economics of Agglomeration*. Chicago: University of Chicago Press and NBER.

Combes, P.P., G. Duranton, L. Gobillon, D. Puga, and S. Roux. 2009. "The Productivity Advantages of Large Cities: Distinguishing Agglomeration from Firm Selection," Working Paper No. 2009-02, Instituto Madrileño de Estudios Avanzados (IMDEA) Ciencias Sociales.

Davis, D. and D.E. Weinstein. 2002. "Bones, Bombs, and Break Points: The Geography of Economic Activity," *American Economic Review*, 92, 1269–1289.

Duranton, G., L. Gobillon, and H.G. Overman. 2007. "Assessing the Effects of Local Taxation Using Microgeographic Data," Working Paper tecipa-247, University of Toronto, Department of Economics.

Eaton, J. and S. Kortum. 2002. "Technology, Geography, and Trade," *Econometrica*, 70, 1741–1779.

Eckstein, Z. and K. Wolpin. 1999. "Why Youths Drop Out of High School: The Impact of Preferences, Opportunities, and Abilities," *Econometrica*, 67, 1295–1340.

Ellison, G., E.L. Glaeser, and W. Kerr. 2007. "What Causes Industry Agglomeration? Evidence from Coagglomeration Patterns," Discussion Paper No. 2133, Harvard Institute of Economic Research.

Epple, D. and H. Sieg. 1999. "Estimating Equilibrium Models of Local Jurisdictions," *Journal of Political Economy*, 107, 645–681.

Glaeser, E.L. and D.C. Mare. 2001. "Cities and Skills," *Journal of Labor Economics*, 19, 316–342.

Gould, E.D. 2007. "Cities, Workers, and Wages: A Structural Analysis of the Urban Wage Premium," *Review of Economic Studies*, 74, 477–506.

Greenstone, M., R. Hornbeck, and E. Moretti. 2008. "Identifying Agglomeration Spillovers: Evidence from Million Dollar Plants," Working Paper No. 13833, National Bureau of Economic Research.

Hahn, J., P. Todd, and W. van der Klaauw. 2001. "Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design," *Econometrica*, 69, 201–209.

Henderson, J.V. 2003. "Marshall's Scale Economies," *Journal of Urban Economics*, 53, 1–28.

Holmes, T.J. 1998. “The Effect of State Policies on the Location of Manufacturing: Evidence from State Borders,” *Journal of Political Economy*, 106, 667–705.

———. 1999. “Localization of Industry and Vertical Disintegration,” *Review of Economics and Statistics*, 81, 314–325.

———. 2005. “The Location of Sales Offices and the Attraction of Cities,” *Journal of Political Economy*, 113, 551–581.

———. 2008. “The Diffusion of Wal-Mart and Economies of Density,” Working Paper No. 13783, National Bureau of Economic Research.

Holmes, T.J. and J.J. Stevens. 2009. “An Alternative Theory of the Plant Size Distribution with an Application to Trade,” manuscript, University of Minnesota.

Hotz, V.J. and R.A. Miller. 1993. “Conditional Choice Probabilities and the Estimation of Dynamic Models,” *Review of Economic Studies*, 60, 497–529.

Keane, M. 2009. “Structural vs. Atheoretic Approaches to Econometrics,” *Journal of Econometrics*, DOI: 10.1016/j.jeconom.2009.09.003.

Keane, M. and K.I. Wolpin. 1997. “The Career Decisions of Young Men,” *Journal of Political Economy*, 105, 473–522.

Kydland, F. and E. Prescott. 1982. “Time to Build and Aggregate Fluctuations,” *Econometrica*, 50, 1345–1370.

Li, B. and Y. Lu. 2009. “Geographic Concentration and Vertical Disintegration: Evidence from China,” *Journal of Urban Economics*, 65, 294–304.

Lucas, R.E. Jr. 1976. “Econometric Policy Evaluation: A Critique,” *Carnegie-Rochester Conference Series on Public Policy*, 1, 19–46.

Lucas, R.E. Jr. and E. Rossi-Hansberg. 2002. “On the Internal Structure of Cities,” *Econometrica*, 70, 1445–1476.

Melitz, M.J. and G.I.P. Ottaviano. 2008. “Market Size, Trade, and Productivity,” *Review of Economic Studies*, 75, 295–316.

Ono, Y. 2003. “Outsourcing Business Services and the Role of Central Administrative Offices,” *Journal of Urban Economics*, 53, 377–395.

Rosenthal, S.S. and W.C. Strange. 2003. “Geography, Industrial Organization, and Agglomeration,” *Review of Economics and Statistics*, 85, 377–393.

———. 2004. “Evidence on the Nature and Sources of Agglomeration Economies,” in J.V. Henderson and J.-F. Thisse (eds.), *Handbook of Urban and Regional Economics*. Amsterdam: Elsevier, pp. 2119–2171.

———. 2008. “The Attenuation of Human Capital Spillovers,” *Journal of Urban Economics*, 64, 373–389.

Shoven, J.B. and J. Whalley. 1984. “Applied General-Equilibrium Models of Taxation and International Trade: An Introduction and Survey,” *Journal of Economic Literature*, 22, 1007–1051.

Sieg, H. 2009. “Structural Local Public Economics,” Paper presented at the conference Regional Science: The Next Fifty Years, Federal Reserve Bank of New York.

Suzuki, J. 2008. “Land Use Regulation as a Barrier to Entry: Evidence from the Texas Lodging Industry,” Working paper, University of Toronto.

van der Klaauw, W. 2002. “Estimating the Effect of Financial Aid Offers on College Enrollment: A Regression-Discontinuity Approach,” *International Economic Review*, 43, 1249–1287.