Manski, C. (1995), Identification Problems in the Social Sciences, Harvard University Press.

6

Simultaneity

A central objective of the social sciences is to learn about the ways in which individuals interact with one another. The enormous body of research on social interactions ranges from economic analysis of the anonymous process by which markets determine prices to ethnographic study of the intensely personal relationships among family members.

The last two chapters of this book examine identification problems that arise when observations of equilibrium outcomes are used to analyze social interactions. Chapter 7 considers the *reflection problem* that arises when a researcher observes the equilibrium distribution of behavior in a population and wishes to learn how the average behavior in some group influences the behavior of the individuals in the group. The present chapter examines the *simultaneity problem* that arises when observations of market transactions are used to study the demand behavior of price-taking consumers or the supply behavior of price-taking (or quantity-taking) firms. Simultaneity also arises in the analysis of nonmarket social interactions, when observations of equilibrium outcomes of games are used to study the reaction functions of the players. So simultaneity is a concern not only of economics but of the social sciences more generally.

6.1. "The" Identification Problem in Econometrics

A classical problem of econometrics is to infer the structure of supply and demand from observations of equilibrium prices and quantities. The basic version of the problem supposes that there is a set of isolated markets for a given product, each market separated from the others in time or in space. Each market has a value for $[s(\cdot), d(\cdot), p, q, x]$. Here x denotes some covariates characterizing a market, q is the quantity of product transacted, and p is the unit price at which these transactions take place. The market demand function $d(\cdot)$ gives the quantity of product that price-taking consumers would purchase if price were set at any level; so $d(t)$ is the quantity demanded if price were set equal to t. The market supply function $s(\cdot)$ gives the quantity of product that price-taking firms would offer if price were set at any level; so $s(t)$ is the quantity supplied if price were set equal to t. The transaction (p, q) is assumed to be an equilibrium outcome; that is, price p makes q both the quantity demanded and the quantity supplied. Formally, (p, q) is assumed to satisfy simultaneously the two conditions

$$(6.1) \quad q = d(p)$$

and

$$(6.2) \quad q = s(p).$$

Markets vary in their values of $[s(\cdot), d(\cdot), p, q, x]$. This heterogeneity is expressed by treating supply, demand, transactions, and covariates as random variables with some distribution $P[s(\cdot), d(\cdot), p, q, x]$. Let $P[s(\cdot), d(\cdot) \mid x]$ denote the distribution of supply and demand functions among markets sharing the same covariates. Econometric analysis seeks to learn about $P[s(\cdot), d(\cdot) \mid x]$ when observations of (p, q, x) are obtained by some sampling process, such as random sampling of markets, that reveals $P(p, q, x)$. Knowledge of $P(p, q, x)$ does not suffice to identify $P[s(\cdot), d(\cdot) \mid x]$. This is the *simultaneity problem.*

Analysis of simultaneity was so central to the early development of econometrics that it was long common for econometricians to think of identification and simultaneity as synonymous. Many econometrics texts, even recent ones, discuss identification only in the context of simultaneity. Particularly revealing is the title chosen by Fisher (1966) for his monograph on the simultaneity problem. He titled the book *The Identification Problem in Econometrics* and justified this choice by writing in the preface (p. vii): "Because the simultaneous equation context is by far the most important one in which the identification problem is encountered, the treatment is restricted to that context." From today's perspective, Fisher's judgment of the preeminence of simultaneity among all identification problems seems strained. Nevertheless, simultaneity remains an important problem of econometrics.

6.2. The Linear Market Model

The problem of learning $P[s(\cdot), d(\cdot) \mid x]$ may be posed for any specification of the covariates x. Given such a specification, the inferences that may be drawn depend on what is known a priori about the structure of supply and demand. Econometric analysis has long centered on the linear model

$$(6.3a) \quad s(t) = \beta_1 t + x'\alpha_1 + u_1$$

$$(6.3b) \quad d(t) = \beta_2 t + x'\alpha_2 + u_2$$

$$(6.3c) \quad E(u_1, u_2 \mid x) = 0.$$

Study of this linear model was initiated in the 1920s and crystallized by the early 1950s (see Hood and Koopmans, 1953).

Equation (6.3a) states that supply $s(\cdot)$ is a linear function of price t, with the same slope β_1 in each market. The supply function varies across markets only in its intercept $x'\alpha_1 + u_1$. This intercept is itself linear in x, a K-dimensional covariate vector observed by the researcher, and in u_1, a scalar covariate expressing determinants of supply that are unobserved by the researcher.

Equation (6.3b) imposes analogous restrictions on demand. Taken together, (6.3a) and (6.3b) transform the problem of learning the distribution of supply and demand into one of learning the parameters $(\beta_1, \beta_2, \alpha_1, \alpha_2)$ and the distribution $P(u_1, u_2 \mid x)$ of unobserved covariates. Condition (6.3c) is prior information about $P(u_1, u_2 \mid x)$. The mean of (u_1, u_2) in markets with covariates x equals zero.

Equations (6.3a) through (6.3c) imply that the mean regressions on x of the supply and demand functions have the linear form

(6.4a) $E[s(t) \mid x] = \beta_1 t + x'\alpha_1$

(6.4b) $E[d(t) \mid x] = \beta_2 t + x'\alpha_2$.

That is, if price were set equal to any value t, then the mean quantity supplied in markets with covariates x would be $\beta_1 t + x'\alpha_1$ and the mean quantity demanded in these markets would be $\beta_2 t + x'\alpha_2$.

The Algebra of Identification

When the linear model (6.3) is combined with the equilibrium conditions (6.1) and (6.2), we find that the equilibrium market outcome is

(6.5a) $q = x'\pi_1 + (\beta_1 u_2 - \beta_2 u_1)/(\beta_1 - \beta_2)$

(6.5b) $p = x'\pi_2 + (u_2 - u_1)/(\beta_1 - \beta_2)$,

where

(6.6a) $\pi_1 \equiv (\beta_1 \alpha_2 - \beta_2 \alpha_1)/(\beta_1 - \beta_2)$

(6.6b) $\pi_2 \equiv (\alpha_2 - \alpha_1)/(\beta_1 - \beta_2)$.

Thus equilibrium quantities and prices are linear functions of (x, u_1, u_2), with parameters that are known functions of the supply and demand parameters $(\beta_1, \beta_2, \alpha_1, \alpha_2)$. Equations (6.5a) and (6.5b) are commonly called the *reduced form* of the linear market model.

It follows from equations (6.5a), (6.5b), and (6.3c) that the mean regressions of equilibrium quantity and price on x are

(6.7a) $E(q \mid x) = x'\pi_1$

(6.7b) $E(p \mid x) = x'\pi_2$.

The sampling process identifies $E(q \mid x)$ and $E(p \mid x)$. Hence the parameters π_1 and π_2 are identified provided only that the covariates x are not perfectly collinear.

The function (6.6) mapping the supply and demand parameters $(\beta_1, \alpha_1, \beta_2, \alpha_2)$ into the reduced-form parameters (π_1, π_2) is many-to-one rather than one-to-one. So identification of (π_1, π_2) does not suffice to identify $(\beta_1, \alpha_1, \beta_2, \alpha_2)$. But the latter parameters are identified if knowledge of (π_1, π_2) is combined with suitable prior restrictions on $(\beta_1, \alpha_1, \beta_2, \alpha_2)$.

By far the most common practice is to invoke exclusion restrictions asserting that some components of α_1 (or α_2) equal zero while the corresponding components of α_2 (or α_1) are nonzero. For example, suppose it is known that $\alpha_{1K} = 0$ but $\alpha_{2K} \neq 0$. That is, holding $x_1, \ldots, x_{K-1}$ fixed, the mean regression of supply on x does not vary with x_K but the mean regression of demand on x does vary with x_K. Then $\pi_{1K}/\pi_{2K} = \beta_1$, so β_1 is identified. Moreover, $\pi_{1k} - \beta_1\pi_{2k} = \alpha_{1k}$, for $k = 1, \ldots, K$, so α_1 is identified. Thus exclusion of a component of α_1 identifies the mean supply regression (6.4a). Similarly, exclusion of a component of α_2 identifies the mean demand regression (6.4b). Observe that these results are obtained without imposing the usual economic assumptions that supply functions are upward sloping (that is, $\beta_1 \geq 0$) and demand functions are downward sloping (that is, $\beta_2 \leq 0$).[1]

Suppose that one can exclude at least one component of α_1 and at least one (different) component of α_2. Then the entire distribution of supply and demand is identified. The exclusion restrictions identify $(\beta_1, \alpha_1, \beta_2, \alpha_2)$, so we need only to identify the distribution $P(u_1, u_2 \mid x)$ of unobserved covariates. Rewrite the reduced-form equations (6.5) in the form

$$(6.5a') \quad \beta_1 u_2 - \beta_2 u_1 = (\beta_1 - \beta_2)(q - x'\pi_1)$$

$$(6.5b') \quad u_2 - u_1 = (\beta_1 - \beta_2)(p - x'\pi_2).$$

It follows that

$$(6.8a) \quad u_1 = (q - x'\pi_1) - \beta_1(p - x'\pi_2)$$

$$(6.8b) \quad u_2 = (q - x'\pi_1) - \beta_2(p - x'\pi_2).$$

With $(\pi_1, \pi_2, \beta_1, \beta_2)$ known, we can determine u_1 and u_2 in each market where (p, q, x) is observed; so we effectively observe realizations of (u_1, u_2, x). Thus $P(u_1, u_2 \mid x)$ is identified.

Predicting the Impact of a Tax

Identification of the distribution of supply and demand is not necessary to predict transactions in markets whose outcomes are observable. The sampling process identifies $P(p, q \mid x)$, so there is no need to learn $P[s(\cdot), d(\cdot) \mid x]$. Economists invoke models of market equilibrium in order to predict market outcomes in new settings. See Marschak (1953).

A typical application is to predict the transactions that would occur if a tax were to be placed on each unit of a product. Suppose that a tax of level τ is introduced. The value of τ may be positive or negative. In the latter case τ is usually called a subsidy, but there is no need to make this semantic distinction.

Let t continue to denote the price received by firms, so $t + \tau$ is now the price paid by consumers. Then equations (6.3a) and (6.3c) of the linear market model remain unchanged, but the demand specification of equation (6.3b) changes to

$$(6.9) \quad d(t) = \beta_2(t + \tau) + x'\alpha_2 + u_2.$$

Inserting (6.3a) and (6.9) into the equilibrium conditions (6.1) and (6.2) now yields not (6.5) but rather

$$(6.10a) \quad q = x'\pi_1 + \beta_1\beta_2\tau/(\beta_1 - \beta_2) + (\beta_1 u_2 - \beta_2 u_1)/(\beta_1 - \beta_2)$$

$$(6.10b) \quad p = x'\pi_2 + \beta_2\tau/(\beta_1 - \beta_2) + (u_2 - u_1)/(\beta_1 - \beta_2).$$

Comparing (6.10) with (6.5) shows that the impact of the tax on market outcomes is to shift the distribution of equilibrium quantities and prices. The mean regression of (q, p) on x, previously given in (6.7), is now

$$(6.11a) \quad E(q \mid x) = x'\pi_1 + \beta_1\beta_2\tau/(\beta_1 - \beta_2)$$

$$(6.11b) \quad E(p \mid x) = x'\pi_2 + \beta_2\tau/(\beta_1 - \beta_2).$$

Thus observation of market outcomes in the absence of a tax does not suffice to predict outcomes following introduction of a tax. But these outcomes can be predicted if the slope parameters β_1 and β_2 of the supply and demand functions are identified.

6.3. Equilibrium in Games

Many social interactions, from divorce proceedings to union-management negotiations to superpower rivalry, can usefully be thought of as games. The basic two-person game imagines that each of two players must select an action. Player 1 chooses an action from some set of feasible choices and player 2 chooses an action from his set of possibilities.

It is common to assume that the players have *reaction functions* $r_1(\cdot)$ and $r_2(\cdot)$ specifying the action that each would choose as a function of the action chosen by the other. Thus $r_1(t_2)$ specifies the action that player 1 would choose if player 2 were to select action t_2. Similarly, $r_2(t_1)$ specifies the action that player 2 would choose if player 1 were to select action t_1. An equilibrium of the game is a pair of mutually consistent actions. That is, (y_1, y_2) is an equilibrium pair of actions if

$$(6.12) \quad y_1 = r_1(y_2)$$

and

$$(6.13) \quad y_2 = r_2(y_1).$$

The problem of interest is to learn the distribution $P[r_1(\cdot), r_2(\cdot) \mid x]$ of reaction functions across games with specified covariates x. The simultaneity problem arises when one attempts to infer $P[r_1(\cdot), r_2(\cdot) \mid x]$ from observations of equilibrium outcomes.

Econometric analysis of two-person games has centered on the linear model

$$(6.14a) \quad r_1(t_2) = \beta_1 t_2 + x'\alpha_1 + u_1$$

(6.14b) $r_2(t_1) = \beta_2 t_1 + x'\alpha_2 + u_2$

(6.14c) $E(u_1, u_2 \mid x) = 0,$

which is analogous to the linear market model (6.3). Conditions (6.12) and (6.13) imply that the equilibrium outcome is

(6.15a) $y_1 = x'\pi_1 + (\beta_1 u_2 + u_1)/(1 - \beta_1\beta_2)$

(6.15b) $y_2 = x'\pi_2 + (\beta_2 u_1 + u_2)/(1 - \beta_1\beta_2),$

where

(6.16a) $\pi_1 \equiv (\beta_1\alpha_2 + \alpha_1)/(1 - \beta_1\beta_2)$

(6.16b) $\pi_2 \equiv (\beta_2\alpha_1 + \alpha_2)/(1 - \beta_1\beta_2).$

Hence the mean regressions of (y_1, y_2) on x are

(6.17a) $E(y_1 \mid x) = x'\pi_1$

(6.17b) $E(y_2 \mid x) = x'\pi_2.$

Analysis of identification follows the same lines as in the linear market model. The sampling process identifies $E(y_1 \mid x)$ and $E(y_2 \mid x)$, so π_1 and π_2 are identified if the covariates x are not perfectly collinear. The function (6.16) mapping $(\beta_1, \alpha_1, \beta_2, \alpha_2)$ into (π_1, π_2) is many-to-one rather than one-to-one, so knowledge of π_1 and π_2 does not identify $(\beta_1, \alpha_1, \beta_2, \alpha_2)$. But the latter parameters are identified if exclusion restrictions can be invoked. For example, suppose it is known that $\alpha_{1K} = 0$ but that $\alpha_{2K} \neq 0$. Then $\pi_{1K}/\pi_{2K} = \beta_1$, so β_1 is identified. Moreover, $\pi_{1k} - \beta_1\pi_{2k} = \alpha_{1k}$, for $k = 1, \ldots, K$, so α_1 is identified.

Ehrlich, the Supreme Court, and the National Research Council

A classical problem in criminology is to learn the deterrent effect of sanctions on criminal behavior. In the early 1970s it became common for criminologists to analyze observed crime rates and sanction levels as equilibrium outcomes of two-person games, wherein criminals (player 1) choose a crime rate and society (player 2) chooses sanctions. Linear reaction functions of the form (6.14) were used to specify the crime rate that criminals would choose if sanctions were set at any level t_2 and the sanctions that society would choose if the crime rate were t_1. In this setting, the parameter β_1 measures the deterrent effect of sanctions, that is, the change in the crime rate that would occur if sanctions were set at different levels.

The simultaneity problem in inference on deterrence became a concern beyond the community of academic criminologists when the Solicitor General of the United States (Bork et al., 1974) argued to the Supreme Court that a study by Isaac Ehrlich provided empirical evidence on the deterrent effect of capital punishment. Ehrlich (1975) used annual data on murders and sanctions in the United States to estimate a "murder supply" function specifying the murder rate that would occur as a function of sanctions levels, including the risk of capital punishment faced by a convicted murderer. He concluded (1975, p. 398): "In fact, the empirical analysis suggests that on the average the tradeoff between the execution of an offender and the lives of potential victims it might have saved was of the order of 1 for 8 for the period 1933–1967 in the United States."

This finding, and its citation before the Supreme Court as evidence in support of capital punishment, generated considerable controversy. A constructive outcome was a series of critiques of the Ehrlich study by Passell and Taylor (1975), Bowers and Pierce (1975), and Klein, Forst, and Filatov (1978), among others. Moreover, a panel of the National Research Council (NRC) was established to investigate in depth the problem of inference on deterrence (Blumstein, Cohen, and Nagin, 1978).

The NRC Panel on Research on Deterrent and Incapacitative Effects focused much of its attention on the simultaneity problem and stressed the difficulty of finding plausible exclusion restrictions to identify deterrent effects. The panel also examined in depth two other identification problems presenting obstacles to inference on deterrence: error in measuring crime rates and confounding of deterrence and incapacitation. In all, the panel report is an exceptionally clearheaded portrayal of the difficulties inherent in the empirical study of deterrence.

Regarding the deterrent effect of capital punishment, the panel concluded (p. 62): "The current evidence on the deterrent effect of capital punishment is inadequate for drawing any substantive conclusion." Cautious scientific assessments of this sort are not unusual in NRC studies, and are usually followed by calls for more research. But this NRC panel went on to draw a more unusual conclusion (p. 63):

> In undertaking research on the deterrent effect of capital punishment, however, it should be recognized that the strong value content associated with decisions regarding capital punishment and the high risk associated with errors of commission make it likely that any policy use of scientific evidence on capital punishment will require extremely severe standards of proof. The nonexperimental research to which the study of the deterrent effects of capital punishment is necessarily limited almost certainly will be unable to meet those standards of proof. Thus, the Panel considers that research on this topic is not likely to produce findings that will or should have much influence on policymakers.

This conclusion is both admirable and distressing. It is admirable that a panel of distinguished social scientists was willing to declare that social science research likely cannot resolve a behavioral question of vital public concern. Powerful incentives induce many researchers to maintain strong assumptions in order to draw strong conclusions. These researchers rejected the temptation.

The conclusion is distressing for the vaccuum it leaves. If research on the deterrent effect of capital punishment "should not have much influence on policymakers," what should influence them? Should policymakers reach conclusions about deterrence based solely on their own observations and reasoning, unaided by research? If so, how should they cope with the inferential problems that stymie social scientists? Should they succumb to the temptation to maintain whatever assumptions are needed to reach conclusions? Or should they give no weight to deterrence and base policy only on their normative views of just punishment? Neither option is appealing.

6.4. Simultaneity with Downward-Sloping Demand

Analysis of the simultaneity problem has long been confined to the linear models described in the preceding sections and to certain extensions thereof. I have found that fresh insights emerge if one puts aside these models and examines the probabilistic structure of the identification problem posed by simultaneity.[2] The analysis in this section focuses on inference on demand, but applies as well to inference on supply and to inference on reaction functions in games.

Simultaneity Is Selection

Simultaneity is actually a problem of censored outcomes. To see this, let x be a specified point on the support of the covariate distribution and let t be a specified price. Consider the distribution $P[d(t) \mid x]$ of the quantity demanded at price t in markets with covariates x. Write

$$(6.18)\quad P[d(t) \mid x] = P[d(t) \mid x, p = t]P(p = t \mid x) + P[d(t) \mid x, p \neq t]P(p \neq t \mid x).$$

Assume that condition (6.1) holds, so the transaction in each market lies on that market's demand function. Then

$$(6.19)\quad P[d(t) \mid x, p = t] = P(q \mid x, p = t).$$

The sampling process reveals $P(p = t \mid x)$, $P(p \neq t \mid x)$, and $P(q \mid x, p = t)$, but does not reveal $P[d(t) \mid x, p \neq t]$. This is precisely the selection problem. The selection probability is $P(p = t \mid x)$. The censoring probability is $P(p \neq t \mid x)$. The distribution of outcomes conditional on selection is $P[d(t) \mid x, p = t]$, or $P(q \mid x, p = t)$ by (6.19). The distribution of outcomes conditional on censoring is $P[d(t) \mid x, p \neq t]$. One wishes to learn about $P[d(t) \mid x]$, the distribution of outcomes that would be observed if price were set equal to t in all markets with covariates x.

Suppose one knows that condition (6.1) holds but has no prior information restricting the structure of demand or supply. Then observations of outcomes in markets where $p \neq t$ reveal nothing about the censored demand distribution $P[d(t) \mid x, p \neq t]$. Thus the *worst-case* analysis of Section 2.2 applies.

In the presence of prior information, observations of outcomes in markets where $p \neq t$ may be informative about the censored demand distribution $P[d(t) \mid x, p \neq t]$. We saw in Section 6.2 that the linear market model accompanied by an exclusion restriction identifies $P[d(t) \mid x]$. Here I consider a different kind of prior information.

Ordered Outcomes

If an economist is willing to assume anything about the structure of demand, it generally is that demand is weakly downward sloping; that is,

$$(6.20) \quad t' > t \Rightarrow d(t') \leq d(t).$$

Yet the literature on simultaneity has not studied the identifying power of this most common economic assumption.

It is intuitive that the assumption of downward-sloping demand should have identifying power. Consider Figure 6.1. If we know only

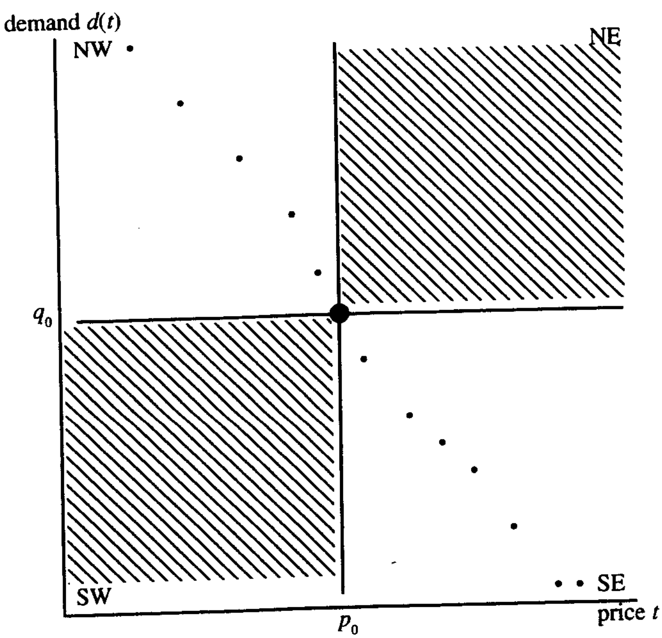


Figure 6.1 Feasible demand functions

that (6.1) holds, then observation of a market outcome (p_o, q_o) reveals only that $d(\cdot)$ is some function passing through the point (p_o, q_o). But if we know that (6.1) and (6.20) hold, then observation of (p_o, q_o) reveals that the downward-sloping $d(\cdot)$ must lie entirely within the northwest (NW) and southeast (SE) regions of the figure, as does the curve drawn.

In Chapter 2, we studied the simplest nontrivial case of assumption (6.20). Recall that the literature analyzing treatment effects supposes that a binary variable z determines which of two outcomes is observed; outcome y_1 is observed if $z = 1$ and outcome y_0 is observed if $z = 0$. In Section 2.6 we examined the identifying power of the *ordered outcomes* assumption[3]

(6.21) $\quad y_1 \le y_0.$

Assumption (6.21) is the special case of (6.20) with $t = 0$ and $t' = 1$. Henceforth I use the term *ordered outcomes* to refer to assumption (6.20) and not just to the special case (6.21).[4]

Bounds on Conditional Probabilities

In Manski (1994c), I show that assumptions (6.1) and (6.20) imply bounds on various features of the distribution $P[d(t) \mid x]$. It is particularly simple to see that these assumptions imply bounds on the probability $P[d(t) \le r \mid x]$ that the quantity demanded at price t is smaller than a specified constant r.

Figure 6.2 decomposes the possible market outcomes into four regions based on the position of (p, q) relative to (t, r). Given that demand is downward sloping, each observation of (p, q) in the SW region implies that $d(t) \le r$ and each observation in the NE region implies that $d(t) > r$. Observations of (p, q) in the SE and NW regions do not reveal whether $d(t)$ is less than or greater than r. Formally, the following holds in each region:[5]

$$\text{SW region:} \quad p \le t \cap q \le r \Rightarrow d(t) \le d(p) = q \le r$$

$$\text{SE region:} \quad p > t \cap q \le r \Rightarrow d(t) \ge d(p) = q$$

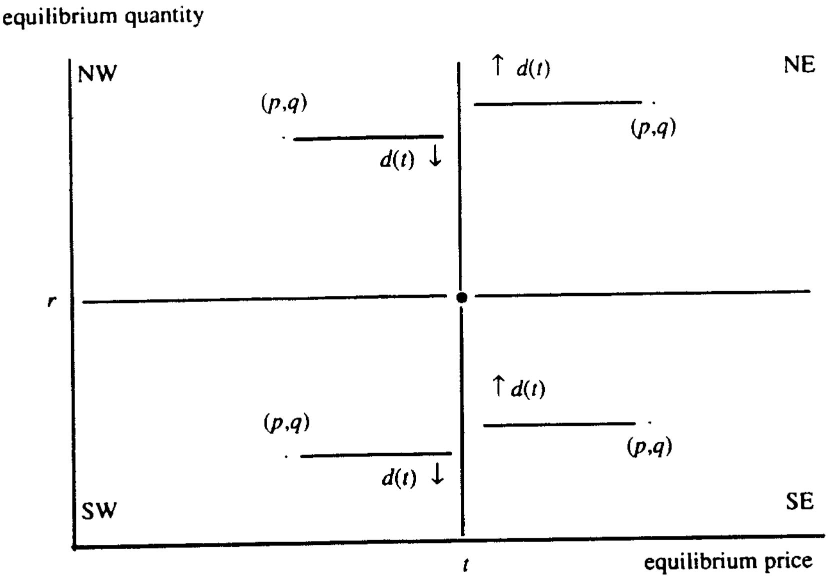


Figure 6.2 Identification of $P[d(t) \leq r \mid x]$.

NW region: $p < t \cap q > r \Rightarrow d(t) \leq d(p) = q$

NE region: $p \geq t \cap q > r \Rightarrow d(t) \geq d(p) = q > r.$

So we have this sharp bound on $P[d(t) \leq r \mid x]$:

(6.22) $$P(p \leq t \cap q \leq r \mid x) \leq P[d(t) \leq r \mid x]$$
$$\leq 1 - P(p \geq t \cap q > r \mid x).$$

The width of the bound depends on the distribution of market outcomes. At one extreme, the support of $P(p, q \mid x)$ may be concentrated in the SE and NW regions of Figure 6.2. Then (6.22) becomes $0 \leq P[d(t) \leq r \mid x] \leq 1$. At the other extreme, the support of $P(p, q \mid x)$ may be concentrated in the SW and NE regions. Then (6.22) becomes $P[d(t) \leq r \mid x] = P(p \leq t \cap q \leq r \mid x)$. So the equilibrium condition (6.1) and the ordered outcome assumption (6.20) may reveal nothing about $P[d(t) \leq r \mid x]$ or may identify this quantity.

The Effect of Policing on Crime

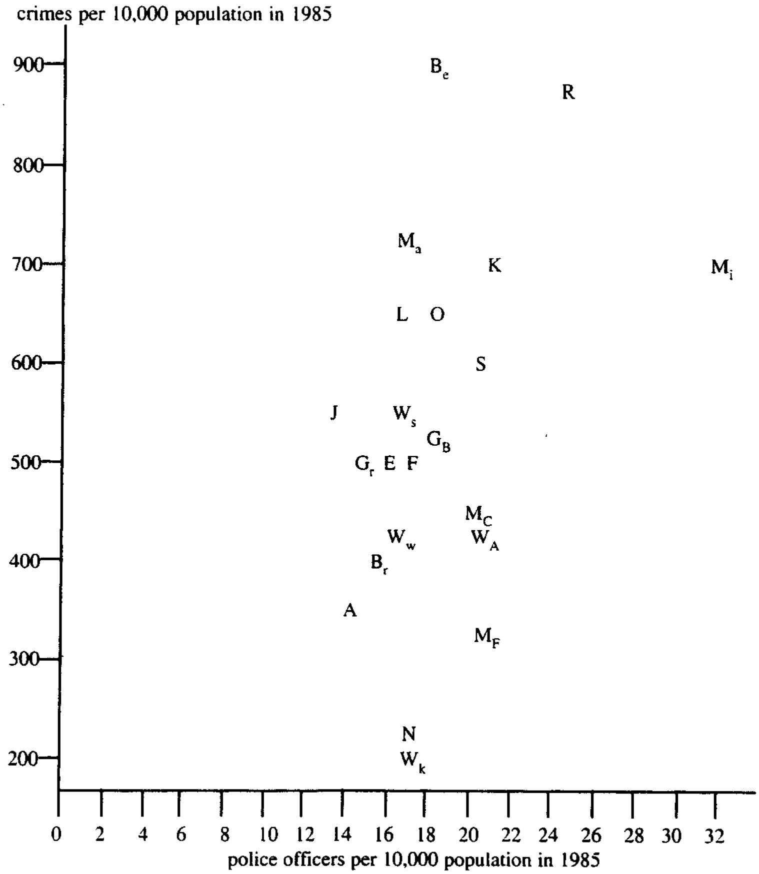
Section 6.3 called attention to the problem of inference on the deterrent effect of sanctions on criminal behavior. I shall use this inferential problem to provide an empirical illustration of the bound (6.22).

The objective is to use observations of crime rates and deterrence policies to study the effect of deterrence policies on criminal behavior. The standard setup assumes a set of isolated jurisdictions. Each jurisdiction has a crime function $d(\cdot)$ giving the crime rate that would occur if deterrence (that is, the price of crime) were set at any level. Some process determines the actual deterrence level p in each jurisdiction. The realized crime rate is then $q = d(p)$. The problem is to learn about the distribution $P[d(\cdot) \mid x]$ of crime functions among jurisdictions with covariates x. Our interest is to learn what inferences are possible if it is assumed only that crime is a weakly decreasing function of the deterrence level.

Empirical studies have measured crime and deterrence in many different ways. To provide an accessible illustration, I use data for American cities with populations over 25,000 published in U.S. Census Bureau (1988b), *County and City Data Book 1988*. The crime rate is an FBI estimate of the number of serious crimes committed per 100,000 resident population in the year 1985 (table C, item 31). The deterrence measure is an FBI estimate of the number of police officers per 10,000 resident population in 1985 (table C, item 33). I use these crime rates and police densities as reported, except that I rescale the crime rate so that it has the same population base as the police density.

I focus on the twenty-two cities in the state of Wisconsin.[6] These observations form a data set large enough to yield interesting findings, but small enough to permit one to follow the calculations easily. Figure 6.3 presents the raw data and displays the configuration of police densities and crime rates.

Assume that the crime functions in these Wisconsin cities are a random sample of size $N = 22$ drawn from the distribution $P[d(\cdot) \mid x]$ of crime functions in jurisdictions sharing specified covariates x.[7] Then consistent estimates of the bound (6.22) may be obtained by replacing the probabilities defining the bound with the corresponding sample frequencies. Table 6.1 presents a set of such estimates. To keep the discussion centered on the problem of identification, I discuss the table as if its entries are the bounds rather than just estimates thereof.



Key

- A - Appleton (14.0, 379.2)
- B_e - Beloit (18.5, 906.0)
- B_r - Brookfield (15.7, 400.3)
- E - Eau Claire (16.0, 493.9)
- F - Fond du Lac (17.2, 487.2)
- G_B - Green Bay (18.2, 540.7)
- G_r - Greenfield (15.1, 508.4)
- J - Janesville (13.4, 561.5)
- K - Kenosha (19.8, 690.4)
- L - La Crosse (17.2, 647.1)
- M_a - Madison (17.2, 722.5)
- M_c - Manitowoc (19.4, 458.2)
- M_F - Menomonee Falls (20.3, 341.1)
- M_i - Milwaukee (32.8, 706.6)
- N - New Berlin (17.0, 222.6)
- O - Oshkosh (17.3, 631.5)
- R - Racine (24.1, 857.2)
- S - Sheboygan (19.2, 601.3)
- W_k - Waukesha (17.1, 221.2)
- W_s - Wausau (16.5, 552.9)
- W_w - Wauwatosa (17.5, 450.8)
- W_A - West Allis (20.2, 425.6)

Figure 6.3 Crime rates and police densities in Wisconsin cities. (Data source: U.S. Bureau of the Census, *County and City Data Book 1988*, table C.)

Table 6.1 Estimated bounds on features of $P[d(\cdot) \mid x]$

	$P[d(t) \leq r \mid x]$			
	$r = 200$	$r = 400$	$r = 600$	$r = 800$
$t = 12$	[.00, .00]	[.00, .18]	[.00, .64]	[.00, .91]
$t = 16$	[.00, .23]	[.05, .36]	[.23, .64]	[.23, .91]
$t = 20$	[.00, .82]	[.14, .86]	[.55, .91]	[.77, .95]
$t = 24$	[.00, .91]	[.18, .91]	[.64, .91]	[.86, .95]

The most striking feature is how much the estimated bounds vary in width as (t, r) varies. We are able to draw only very weak inferences when $(t, r) = (12, 800)$ and $(t, r) = (24, 200)$. In these cases, the table entries are $P[d(12) \leq 800 \mid x] \in [.00, .91]$ and, likewise, $P[d(24) \leq 200 \mid x] \in [.00, .91]$. We can, however, draw rather strong inferences when $(t, r) = (12, 200)$ and $(t, r) = (24, 800)$. Here $P[d(12) \leq 200 \mid x] = 0$ and $P[d(24) \leq 800 \mid x] \in [.86, .95]$. That is, taking the estimates at face value, we find that setting the police density equal to 12 officers per 10,000 population always yields a crime rate larger than 200 crimes per 10,000 population. Setting the police density equal to 24 officers per 10,000 population has at least a .86 chance of yielding a crime rate less than or equal to 800 crimes per 10,000 population.