Physician-induced demand for childbirths

David Dranove* and Paul Wehner

Department of Management and Strategy, Kellogg Graduate School of Management, Northwestern University, 2001 Sheridan Road, Evanston, IL 60208, USA

Received April 1993, final version received July 1993

Abstract

A controversial technique for testing the hypothesis that physicians induce demand involves two stage least squares (TSLS) regression analysis of cross-section data on physician supply and utilization. This paper tests the power of TSLS by applying it where there is at most only a trivial amount of demand inducement – the demand for childbirths. We find 'evidence' of inducement of childbirths, calling into question the validity of the TSLS approach. This unlikely finding may be traced to at least two factors: The first stage regression is not identified and the second stage regression does not adequately address border crossing.

Key words: Demand inducement; Physician supply

JEL classification: I11, C52

1. Introduction

The notion that suppliers of medical care can create demand for their services dates at least to Roemer (1961), who posited that a hospital could fill its beds regardless of the underlying demand for hospital care. Fuchs' (1978) model crystallized the inducement hypothesis. In Fuchs' model, which resembles simple models of advertising and quality choice, inducement shifts the demand curve. A key difference between Fuchs' model and models of advertising and quality choice is that inducement appears to be costless in Fuchs' model. Dranove (1988) suggested that physicians induce demand by recommending services that do not pass objective patient benefit–cost thresholds. While these models differ in substance, they all agree that

*Corresponding author. Fax 708-467-7777.

1 Fuchs' model only makes sense if physicians are not profit maximizers. See Dranove (1988) for a further discussion.

0167-6296/94/$07.00 © 1994 Elsevier Science B.V. All rights reserved
SSDI 0167-6296(93)E0048-T
suppliers of medical services have the ability to influence the demand for their products. Indeed, we suspect that there are few health economists who question the theoretical merits of the inducement hypothesis.

The inducement hypothesis has significant policy implications. For example, it is often claimed that physicians who own diagnostic imaging equipment are more likely to order tests than are physicians who do not own equipment (Crane, 1992; Hillman et al., 1992; Mitchell and Scott, 1992a). As a result, many states have restricted self-referrals by physicians who own imaging equipment. To take another example, the Health Care Finance Administration (HCFA) was concerned that surgeons might respond to the Resource Based Relative Value Scale by inducing demand for surgery. HCFA seriously considered but ultimately did not impose ceilings on the total amount they would pay for surgeries nationwide. As a final example, it is often claimed that the abundance of surgical specialists in the United States is an important contributor to our relatively high utilization of surgical procedures, and call for reductions in specialty surgical residency programs (see, for example, Schroeder, 1992).

The inducement theory may be sound, and the policy implications clear, but the theory is not wholly convincing and the evidence on the actual magnitude of demand inducement is controversial. One popular technique for testing the inducement hypothesis is to use two-stage least squares (TSLS) with cross-section data on physician supply and utilization. This technique has received much criticism on econometric grounds (see Auster and Oaxaca, 1981; Feldman and Sloan, 1988; Phelps, 1986; Ramsey, 1981). We test the power of the TSLS approach by examining inducement in a situation where it almost surely does not exist. We find that in simple specifications that fail to adequately measure critical demand variables, TSLS leads to the absurd conclusion that obstetricians induce demand for childbirths. We identify at least two reasons for this result: (1) The failure to empirically identify the first stage regression; (2) The failure to account for border crossing. We conclude that TSLS as currently practised should be abandoned.

2. Reviewing the theory and evidence on inducement

Inducement theory is motivated by the long-standing cross-section correlation between the supply of medical services on the one hand and price and consumption on the other. The seminal model of inducement was offered by Fuchs (1978). In his model, physicians induce demand by undertaking activities that shift out the demand curve. Fuchs' model raises an obvious question: If physicians are reimbursed on a fee for service basis, why do they not induce to the maximum? There are a number of possible explanations.
First, physicians may be satisficers, inducing sufficient demand to reach a target level of income. Second, inducement may be costly in that physicians feel 'guilt' that they trade off against the financial benefits. Lastly, the market may constrain physician ability to induce – patients who suspect that they are victims of inducement may take their business elsewhere.

Satterthwaite (1982) offers an alternative explanation for the correlation between supply and price. He suggests that some markets may be naturally less 'competitive' than others, due to variations in local regulations, or variation in consumer shopping behavior. Less competitive markets will have higher price-cost margins, and therefore attract more physicians. Physician migration to high price/cost margin communities will 'induce' the positive correlation between supply and price. While Satterthwaite does not discuss the correlation between supply and quantity, some possibilities are suggested by his analysis. Increased MD supply could reduce the time price of visiting a physician. Also, increased MD supply could lead to more intense quality competition, thereby leading to increased consumption.

Evidence supporting the inducement hypothesis takes two forms. The first form compares health services utilization when physicians are compensated on a fee for service basis with utilization when physicians are capitated or salaried. It is well accepted that capitated and salaried physicians provide less medical care than physicians paid fee for service. Recent papers find evidence of more subtle links between physician incentives and utilization. These findings have immense policy implications. For example, as a result of work by Crane (1992), Hillman et al. (1992) and Mitchell and Scott (1992a,b) relating physician ordering of diagnostic tests to their ownership of testing equipment, several states have enacted legislation to limit physician 'self-referrals' for diagnostic services. It should be pointed out that most of the studies linking incentives to performance do not attempt to sort out causality. Although it is presumed that incentives alter performance, it is possible that physicians with a propensity towards a particular style of practice may seek out a particular type of incentive scheme. Arnould and Debrock (1993) use two stage least squares methods to try to resolve some of these issues. They fail to confirm the earlier literature's presumption that incentives do matter, although their underlying model may be inappropriate for the question in hand.

The second form of evidence on inducement identifies the relationship between the demand for physician services and the supply of physicians. Underlying these studies is the assumption that even if all physicians are paid on a fee-for-service basis, they do not always induce demand to the maximum extent possible. Although this literature tends to be more widely cited on the subject of inducement, it actually conducts a weaker test. To see

---

2 See Gaynor and Gertler (1992) for a review of this literature.
why, suppose that these studies found no link between supply and quantity. Could we conclude that physicians are not inducing demand in Fuchs' sense? No! Physicians paid on a fee-for-service basis may still be recommending unnecessary medical care, but their inducement efforts might be identical across all market areas. Hence, we would not observe any cross-sectional correlation between supply and inducement activity.

Fuchs (1978) is the seminal example of the weak test of inducement. Cromwell and Mitchell (1986) provide important modifications to the basic Fuchs approach. Both papers use TSLS on regional cross-sections. TSLS is necessary because causality between physician supply and utilization can run both ways. The first stage identifies a physician supply equation and ideally includes variables that are plausibly related to physician supply but not demand. Fuchs' identifying variables appear to be hotel receipts per capita and status as a metropolitan area. Cromwell and Mitchell add a measure of the temperature, and the number of local colleges. Their arguments appear to be that physicians are attracted to 'touristy' locations with nice weather, and that these variables do not instrument for any omitted demand side variables. These authors do not test whether metropoli-

tanness, hotels and temperature are identifiers. Indeed, Auster and Oaxaca (1981) and Feldman and Sloan (1988) question whether the demand equation in traditional TSLS can ever be identified.

A number of recent studies of the effect of physician supply on inducement examine patient level data. A patient-level study by Wilensky and Rossiter (1983) use TSLS and finds only a small amount of inducement— they estimate an elasticity of 10% between MD supply and frequency of 'induced' follow-up visits to MD offices. (Cromwell and Mitchell obtain a similar elasticity for surgeries; Fuchs' estimated elasticity is more than twice these estimates.) Feldman and Sloan (1988) point out that the use of patient level data does not by itself eliminate the need to worry about identification of the first stage equation. They point out that patients in a given region may have a characteristic in common (at least probabilistically) that is related to demand. Absent good first stage identifiers, the second stage coefficient on predicted physician supply will still be biased.

Feldman and Sloan (1988) recommend estimating fixed region effects models in a time-series cross-section in order to eliminate concerns about whether physician supply is correlated with omitted demand. Based on this idea, Jones and Salkever (1993) examine the correlation between short term fluctuations in reported physician supply and utilization within regions. They find that intra-region fluctuations in supply and utilization are uncorrelated,

---

3These papers do not address the possibility that increased physician supply reduces the time price of obtaining services, so that causation runs from supply to use but there is no inducement in the traditional sense.
and conclude that there is no inducement. The fixed effects methodology is not without its own set of potential problems, however. A key unaddressed issue is the endogeneity of intra-region fluctuations in MD supply.

Phelps and Newhouse (1974), Acton (1975) and Escarce (1992) offer another explanation for the observed correlation between supply and utilization. They suggest that increases in physician supply facilitate access. Specifically, they argue that increasing the supply reduces the effective price of care to patients, including time price. Escarce provides some evidence to support this contention. Related to this idea, patients may cross borders to areas with higher physician supplies, particularly if increased supplies are associated with higher quality due to specialization, (Baumgardner, 1988) or quality competition (Dranove and Satterthwaite, 1992).

In spite of these recent challenges, the works of Fuchs, Cromwell and Mitchell, and others continue to be cited as proof of the existence of inducement (e.g. Folland et al., 1993). One possible reason why many researchers and policy analysts persist in believing the results of the TSLS estimates is that challenges to TSLS remain grounded in econometric theory rather than evidence. This paper directly assesses the ability of TSLS methods to correctly sort out uncorrelated demand from true inducement effects. We implement an 'anti-test' – we apply TSLS to a situation where the conclusion about whether inducement exists is foregone, and determine if TSLS is powerful enough to reject the inducement hypothesis in this situation.

3. The 'anti-test' of TSLS

We propose to evaluate the power of TSLS test for inducement by implementing it in a situation where inducement almost surely does not exist. If TSLS cannot reject inducement in a situation in which it does not exist, then it lacks power as a technique for identifying inducement when it does exist. To implement this 'anti-test', we require a type of medical utilization with two characteristics:

1. The supply of physicians is positively correlated with utilization in the cross-section.
2. The argument that the correlation runs from physician supply to utilization lacks face validity.

A type of utilization that fits these characteristics is childbirths. The first characteristic is easily met: The county-level cross-section correlation between the number of childbirths per capita and the number of obstetricians per capita is 0.33, which is significant at $p < 0.01$. The second characteristic also

appears to be met. While obstetricians may influence where and how children are born, it is implausible to think that they can substantially influence the number of childbirths. It is possible that obstetricians in areas of abundant supply may be more likely to prescribe fertility drugs as a way of boosting demand, but we suspect that this effect is of trivial magnitude. Obstetricians may affect the number of childbirths in other ways, such as by providing prenatal care, but this would be an example of either an availability or quality effect, not demand inducement per se. We conclude that if we observe a positive influence of obstetrician supply on childbirths, we would not want to conclude that this was evidence of inducement. Rather, the correlation results from misspecification of the estimation equation, patient migration, or reduced time cost/increased quality.

Our estimation procedure proceeds as follows. First, we estimate TSLS using a set of variables that frequently appear in other inducement studies. These are listed below. The variables JANUARY TEMP, HOTEL RECEIPTS and URBAN are similar to variables used by Fuchs and Cromwell and Mitchell to identify their physician supply equations, and the remaining variables (or variants thereof) appear in Cromwell and Mitchell's second stage regressions. Due to the variety of data sources available to us, the years for our variables do not match up. All variables are calculated at the level of the county. The variables CHILDBIRTHS, OBGYN, SCHOOL and INCOME were obtained from the 1990 Area Resource File. The remaining variables come from USA Counties (1992), which reports a wide variety of information at the county level.

CHILDBIRTHS: The number of childbirths per 1000 capita in 1988.

OBGYN: The number of physicians self-reporting their specialty as obstetrics/gynecology per 1000 capita in 1990.

DIVORCE: Proportion of females who have ever divorced, as of 1980.

INCOME: 1990 per capita income.


SCHOOL: Median highest school year attained, all individuals over 24 years of age, in the year 1988.

HOSPITAL BEDS: Number of hospital beds per 1000 total population in 1985.

HOTEL RECEIPTS: Hotel receipts per capita in 1987.


URBAN: Proportion of county's population residing in urban areas.

WAGE: Average hourly retail wage in 1987.

We considered using a variable indicating racial composition but this was not reported for many counties in our data base. When we excluded those counties and included the measure of race, the results did not change.

The baseline county population is for 1985.
Table 1
Summary statistics: county level data

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std. dev.</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHILDBIRTHS</td>
<td>14.37</td>
<td>3.41</td>
<td>4.52</td>
<td>48.48</td>
</tr>
<tr>
<td>OBGYN</td>
<td>0.045</td>
<td>0.064</td>
<td>0</td>
<td>0.60</td>
</tr>
<tr>
<td>PREDICTED OBGYN</td>
<td>0.045</td>
<td>0.038</td>
<td>-0.031</td>
<td>0.24</td>
</tr>
<tr>
<td>DIVORCE</td>
<td>5.22</td>
<td>1.85</td>
<td>0</td>
<td>15</td>
</tr>
<tr>
<td>FEMALE</td>
<td>0.47</td>
<td>0.06</td>
<td>0.24</td>
<td>0.90</td>
</tr>
<tr>
<td>HOSPITAL BEDS</td>
<td>5.02</td>
<td>5.72</td>
<td>0</td>
<td>72.50</td>
</tr>
<tr>
<td>HOTEL RECEIPTS</td>
<td>1.16</td>
<td>5.02</td>
<td>0</td>
<td>154.83</td>
</tr>
<tr>
<td>INCOME</td>
<td>15.21</td>
<td>3.52</td>
<td>6</td>
<td>39</td>
</tr>
<tr>
<td>JANUARY TEMP</td>
<td>32.82</td>
<td>12.14</td>
<td>0</td>
<td>67</td>
</tr>
<tr>
<td>OVER 65</td>
<td>0.19</td>
<td>0.05</td>
<td>0.03</td>
<td>0.38</td>
</tr>
<tr>
<td>SCHOOL</td>
<td>11.95</td>
<td>0.88</td>
<td>7</td>
<td>16</td>
</tr>
<tr>
<td>UNEMPLOYMENT</td>
<td>6.75</td>
<td>3.34</td>
<td>0.60</td>
<td>33.40</td>
</tr>
<tr>
<td>URBAN</td>
<td>0.36</td>
<td>0.29</td>
<td>0</td>
<td>1.0</td>
</tr>
<tr>
<td>WAGE</td>
<td>8.70</td>
<td>1.23</td>
<td>3.60</td>
<td>15.65</td>
</tr>
<tr>
<td>WHITE</td>
<td>0.48</td>
<td>0.45</td>
<td>0</td>
<td>1.0</td>
</tr>
<tr>
<td>WHITE COLLAR</td>
<td>0.25</td>
<td>0.07</td>
<td>0.03</td>
<td>0.77</td>
</tr>
</tbody>
</table>

WHITE COLLAR: Proportion of county's workforce in professional, specialty or managerial occupations, in 1980.

After estimating TSLS with these variables, we include several demographic variables that one might find in a fertility study:

FEMALE: Percentage of the population that is female in 1988.

OVER65: Percentage of the female population that is over the age of 65 in 1988.

WHITE: Percentage of the population that is white in 1988.

These variables represent 'predisposing health indicators' that are missing in most inducement studies. For example, neither Fuchs nor Cromwell and Mitchell include indicators of predisposing need for surgery.

Table 1 reports statistics for these variables.

This study suffers from a number of specification problems that are common to other inducement studies, including Fuchs, Cromwell and Mitchell, Wilensky and Rossiter, and Jones and Salkever. First, physicians can usually produce more services than just those under consideration in the inducement study. For example, surgeons can provide follow-up care, diagnostic services, and non-surgical care. In our case, obstetricians can provide prenatal care as well as general gynecological services. Thus, in any inducement study, cross-section variation in physician supply or utilization may reflect substitution of physician time across a number of services. Second, there are other sellers who may provide the services provided by physicians in question. For example, primary care practitioners may perform
Table 2
County level regression results: ordinary least squares
variable = CHILDBIRTHS

<table>
<thead>
<tr>
<th>Predictor</th>
<th>(1) Coeff.</th>
<th>t-ratio</th>
<th>(2) Coeff.</th>
<th>t-ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>OBGYN</td>
<td>14.80</td>
<td>13.3</td>
<td>4.70</td>
<td>5.09</td>
</tr>
<tr>
<td>DIVORCE</td>
<td>0.19</td>
<td>5.1</td>
<td>-0.013</td>
<td>-0.44</td>
</tr>
<tr>
<td>HOSPITAL BEDS</td>
<td>-0.052</td>
<td>-5.3</td>
<td>0.050</td>
<td>6.35</td>
</tr>
<tr>
<td>INCOME</td>
<td>-0.15</td>
<td>-6.8</td>
<td>-0.009</td>
<td>-0.52</td>
</tr>
<tr>
<td>SCHOOL</td>
<td>0.025</td>
<td>0.30</td>
<td>0.225</td>
<td>3.49</td>
</tr>
<tr>
<td>UNEMPLOYMENT</td>
<td>-0.011</td>
<td>-0.57</td>
<td>0.045</td>
<td>2.75</td>
</tr>
<tr>
<td>WAGE</td>
<td>0.40</td>
<td>7.6</td>
<td>0.053</td>
<td>1.28</td>
</tr>
<tr>
<td>WHITE COLLAR</td>
<td>3.00</td>
<td>2.54</td>
<td>-3.74</td>
<td>-4.04</td>
</tr>
<tr>
<td>FEMALE</td>
<td>26.01</td>
<td></td>
<td>29.83</td>
<td></td>
</tr>
<tr>
<td>OVER 65</td>
<td>-35.37</td>
<td></td>
<td>-35.62</td>
<td></td>
</tr>
<tr>
<td>WHITE</td>
<td>-0.77</td>
<td></td>
<td>-6.79</td>
<td></td>
</tr>
<tr>
<td>CONSTANT</td>
<td>10.85</td>
<td>10.66</td>
<td>6.26</td>
<td>6.62</td>
</tr>
</tbody>
</table>

Adjusted $R^2$ 0.175 0.516
N 3058 3058

some minor surgery. In our case, midwives substitute for obstetricians in
some markets. Third, patients often receive services outside of their local
community. Demand estimates may be confounded by this border crossing.7

We lack the data necessary to determine the importance of the first two
specification problems. We examine the border crossing problem by reesti-
mating the model using state level data. Although the sharply reduced
sample size makes it difficult to interpret changes in significance levels, any
reduction in the estimated magnitude of the inducement effect would be
consistent with a border crossing problem.

4. Results

Table 2 reports ordinary least squares regression results in which we
include OBGYN directly as a predictor variable. Model (1) reports results
without fertility-related demographic measures; they are included in model
(2). As expected, OBGYN is positively related to CHILDBIRTHS. Based on
the estimates in model (2), we calculate that a 10% increase in the supply of
OBGYN in a county is associated with a 0.15% increase in CHILDBIRTHS.

7 Fuchs is probably not affected by border crossing, since he examines fourteen major census
areas.
Table 3
County level regression results: TSLS first stage regression dependent variables = OBGYN

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Coeff.</th>
<th>t-ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>HOTEL RECEIPTS</td>
<td>0.00080</td>
<td>4.2</td>
</tr>
<tr>
<td>INCOME</td>
<td>0.0039</td>
<td>12.2</td>
</tr>
<tr>
<td>JANUARY TEMP</td>
<td>0.00051</td>
<td>5.9</td>
</tr>
<tr>
<td>SCHOOL</td>
<td>0.0030</td>
<td>2.2</td>
</tr>
<tr>
<td>URBAN</td>
<td>0.098</td>
<td>26.7</td>
</tr>
<tr>
<td>CONSTANT</td>
<td>-0.104</td>
<td>-6.5</td>
</tr>
</tbody>
</table>

Adjusted $R^2$ 0.341

N 3072

or an elasticity of 1.5% at the means. To put it another way, a one standard deviation increase in OBGYN is associated with a 2.1% increase in CHILDBIRTHS.

Table 3 reports the results of our first-stage estimates of OBGYN. All of our ‘identifying’ variables, JANTEMP, HOTELS and URBAN have the expected signs and are significant.8

Table 4 reports the results of our second-stage estimates of CHILDBIRTHS. Model (1) excludes fertility-related demographic variables. These are included in model (2). The coefficient on predicted OBGYN is positive and significant. The coefficient in model (2) implies an elasticity of OBGYN at the means of 8%, which is comparable to the inducement elasticities found in Cromwell and Mitchell and Wilensky and Rossiter. To put it another way, a one standard deviation increase in predicted OBGYN is associated with a 7% increase in CHILDBIRTHS.9

The signs on the control variables in model (1) on Tables 2 and 3 are somewhat inconsistent with expectations. For example, higher levels of DIVORCE should be associated with lower CHILDBIRTHS, and higher supply of HOSPITAL BEDS should have a positive availability effect on the CHILDBIRTHS. In addition, if WAGE represents the opportunity cost of

---

8 We estimated the model using each identifying variable independently. Each was significant in an OLS model at $p < 0.10$ or better. Thus, the results are not an artifact of the choice of any one particular identifier. Of the three identifiers, URBANPOP had the strongest effect (i.e., the highest fertility elasticity). As an anonymous referee pointed out, such a variable is a questionable identifier, since URBANPOP may reflect cultural or ethnic variation, or could indicate referral centers. Both Fuchs and Cromwell and Mitchell use an identifier similar to URBANPOP, so this concern extends to their estimates as well.

9 We obtain similar results if we limit the sample to inlier counties – those with BIRTHPOP not in the upper or lower tails. The second stage regressions are not heteroscedastic with respect to county population. There is no need to test for heteroscedasticity in the first stage since predicted OBGYN is unbiased even in the presence of heteroscedasticity.
Table 4
County level regression results: TSLS second stage regression dependent variable = CHILDBIRTHS

<table>
<thead>
<tr>
<th>Predictor</th>
<th>(1) Coeff.</th>
<th>t-ratio</th>
<th>(2) Coeff.</th>
<th>t-ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>PREDICTED OBGYN</td>
<td>27.53</td>
<td>11.2</td>
<td>26.45</td>
<td>14.0</td>
</tr>
<tr>
<td>DIVORCE</td>
<td>0.13</td>
<td>3.2</td>
<td>-0.15</td>
<td>-4.8</td>
</tr>
<tr>
<td>HOSPITAL BEDS</td>
<td>-0.047</td>
<td>-4.7</td>
<td>0.049</td>
<td>6.4</td>
</tr>
<tr>
<td>INCOME</td>
<td>-0.24</td>
<td>-9.7</td>
<td>-0.12</td>
<td>-6.1</td>
</tr>
<tr>
<td>SCHOOL</td>
<td>-0.13</td>
<td>-1.5</td>
<td>0.17</td>
<td>2.7</td>
</tr>
<tr>
<td>UNEMPLOYMENT</td>
<td>-0.024</td>
<td>-1.1</td>
<td>0.046</td>
<td>2.9</td>
</tr>
<tr>
<td>WAGE</td>
<td>0.406</td>
<td>7.7</td>
<td>0.011</td>
<td>0.26</td>
</tr>
<tr>
<td>WHITE COLLAR</td>
<td>5.25</td>
<td>4.6</td>
<td>-4.83</td>
<td>-5.5</td>
</tr>
<tr>
<td>FEMALE</td>
<td>27.83</td>
<td>34.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>OVER 65</td>
<td>-34.98</td>
<td>-36.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WHITE</td>
<td>-1.07</td>
<td>-9.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CONSTANT</td>
<td>13.14</td>
<td>12.2</td>
<td>8.14</td>
<td>8.8</td>
</tr>
</tbody>
</table>

Adjusted $R^2$          | 0.161      | 0.54    |
Number of observations   | 3058       | 3058    |

Time, it should have a negative coefficient. We see from model (2) in both Tables that the odd signs on the control variables were a result of omitted variable bias. When we include fertility-related demographic variables, the $R^2$ increases substantially, and the signs on most control variables fall in line with expectations.

What is inducing the observed inducement effect? One possibility is that OBGYN supply equation is not identified; i.e., the identifier variables are demand proxies rather than supply proxies. There is a simple test for this possibility. The system of equations that we have estimated may be written as follows:

$$OBGYN = f(CHILDBIRTHS, X) \quad (1)$$

$$CHILDBIRTHS = g(OBGYN, Y) \quad (2)$$

where $X$ and $Y$ represent supply and demand shifters, respectively. For Eq. (2) to be identified, there must be some elements of $X$ that are not in $Y$.

We test for identification by using ordinary least squares to estimate Eq. (2), restricting the set of $Y$ variables to the exogenous predictors used in the second stage regressions discussed above. We then compute the correlations between the residuals from these OLS estimates and the identifying variables HOTEL RECEIPTS, JANUARY TEMP, and URBAN. The respective correlations are 0.0280, 0.0551, and 0.1606. The latter two correlations are
significant at \( p < 0.01 \). Thus, the parameters used to identify Eq. (1) belong in the vector \( Y \) of parameters that shift \( \text{CHILDBIRTHS} \) in Eq. (2), and Eq. (2) is therefore not identified.\(^{10}\)

Another possibility for the observed inducement effect is that the estimates are biased due to border crossing. Markets with high scores on \( \text{URBAN}, \ \text{JANUARY TEMP} \) and \( \text{HOTEL RECEIPTS} \) may in fact be attractive to physicians. As a result, patients travel to those locations to receive care. The second stage estimates thus reflect border crossing rather than inducement. One way to test for this would be to obtain data on where children were born, rather than where the mother resided. Lacking this data, we test for the possible effects of border crossing by aggregating the data to the state level. Results of the state level second stage estimates appear in Table 5.\(^{11}\)

Model (2) in Table 5, which includes fertility-related demographic variables, reports a coefficient on predicted \( \text{OBGYN} \) that is dramatically smaller than those reported above, and is no longer significant. The implied elasticity at the means is now only 5\%, compared with 7\% reported above. These findings all suggest that the TSLS results at the county level may reflect some border crossing, with the effects of \( \text{OBGYN} \) mitigated when the data is aggregated to larger geographic areas.

---

\(^10\) Consistent with these findings, when \( \text{HOTELRECT} \) is used as the only identifier – i.e., it is directly included as an instrument in Eq. (2) – it is not significant at conventional levels.

\(^{11}\) Summary statistics and first stage estimates are available from the authors on request.
5. Discussion

This paper applies TSLS to cross-section data on childbirths and obstetrician supply. While these variables are positively correlated in the cross-section, this correlation is almost surely not an indication of inducement. We use traditional first stage identifying variables – temperature, hotel receipts and urban status – to try to sort out causality. We find that with these ‘identifiers’, TSLS is not sensitive enough to reject the inducement hypothesis. A principle reason is that the chosen ‘identifiers’ do not identify the first stage regression. We question the validity of TSLS estimates of inducement that rely on these identifiers to identify the first stage physician location regression. We recommend that any other suggested identifiers be put to the ‘anti-test’ used in this paper.

References

Arnould, R. and L. DeBrock, 1993, Strategic choices of internal compensation arrangements in HMOs, Mimeo (University of Illinois).
Escarce, J., 1992, Explaining the association between the surgeon supply and utilization, Inquiry 29, 403–415.
Satterthwaite, M., 1982, Competition and equilibrium as a driving force in the health services sector, Mimeo (Northwestern University).