ARTICLES

PUBLIC INFRASTRUCTURE AND EXTERNALITIES IN U.S. MANUFACTURING: EVIDENCE FROM THE PRICE-AUGMENTING AIM COST FUNCTION

GUOHUA FENG

University of North Texas

APOSTOLOS SERLETIS

University of Calgary

In this paper, we propose a price-augmenting asymptotically ideal model (AIM) cost function to investigate the effects of public infrastructure on the performance of the U.S. manufacturing industry, using KLEMS data over the period from 1953 to 2001. In doing so, we make a distinction between the productivity effect and the production factor effect of public infrastructure. This distinction allows us to focus on the more interesting productivity effect by incorporating public infrastructure into the AIM cost function through the efficiency index. Moreover, we specify the growth rate of the efficiency index as a Box–Cox function of public infrastructure and a time trend, a proxy for other technology. The excellent flexibility of our price-augmenting AIM cost function offers many insights regarding the effects of infrastructure on the U.S. manufacturing sector.

Keywords: Flexible Functional Forms, Asymptotically Ideal Model, Nonlinear Cost Function, Efficiency Index, Public Infrastructure

1. INTRODUCTION

The size and significance of the effects of public infrastructure on the economic performance of the private sector have been a hotly debated topic since Aschauer (1989) raised the issue of productivity of infrastructure capital—see, for example, surveys in Gramlich (1994), Sturm et al. (1998), Romp and de Haan (2007), and Hashimzade and Myles (2010). Early studies in this literature find a close correlation between reductions in public capital investment and declining private-sector productivity in the United States and many other developed economies. Most

We would like to thank the Editor, William A. Barnett, and two anonymous referees for comments that greatly improved the paper. Guohua Feng also gratefully acknowledges financial support from the Australian Research Council. Address correspondence to: Apostolos Serletis, Department of Economics, University of Calgary, Calgary, Alberta, T2N 1N4, Canada; e-mail: Serletis@ucalgary.ca.

of the subsequent studies have been aimed at reaching consensus on the extent of these effects. Despite the voluminous literature, however, the issue remains contentious.

The literature investigating the effects of public infrastructure on the productivity of the private sector is dominated by two approaches—the production function approach and the cost function approach. Introduced by Aschauer (1989), the basic idea of the production function approach is to expand an aggregate production function by specifying public infrastructure as a separate factor input. Early studies along this line [see, for example, Aschauer (1989) and Munnell (1990)] find that public investment has a much greater return to private-sector economic performance than does private capital investment. For example, Aschauer (1989) finds the elasticity of output with respect to public capital to range from 0.39 to 0.56, and the marginal product of public capital implied by this result is 100% or more—see Gramlich (1994).

These findings imply that policy measures designed to augment public infrastructure investment could dramatically enhance productivity. In questioning the robustness of the empirical results, subsequent studies using the production function approach focus more on refining the econometric structures by incorporating state and time fixed effects and/or by solving potential econometric problems, such as nonstationarity, spurious correlation, and endogeneity. The estimated effects of public infrastructure investment on private sector productivity vary across these later studies. In particular, some studies obtain estimates similar to the earlier ones [see, for example, Munnell (1990)]; some find estimates with reduced magnitude and significance [see, for example, Tatom (1991) and Kemmerling and Stephan (2002)]; and others even find that the estimated productivity effects disappear [see, for example, Hulten and Schwab (1991), Evans and Karras (1994), and Holtz-Eakin (1994)].

Although less frequently used than the production function approach, the cost function approach offers a different perspective on the effects of public infrastructure investment on the performance of the private sector. Pioneered by Berndt and Hansson (1992), Nadiri and Mamuneas (1994), and Morrison and Schwartz (1996), the cost function approach, in assuming that firms minimize cost subject to a given level of output, treats public infrastructure as an unpaid fixed input. Compared with the production function approach, the cost function approach has some advantages-see Berndt and Hansson (1992) for a detailed discussion of the merits of the cost function approach. It is less likely to suffer from the endogeneity problem, because input prices used in the cost function approach are more likely to be exogenous than input quantities (in this study we first assume that input prices are exogenous and then conduct a robustness test using a two-stage least squares approach); it usually employs flexible functional forms in place of the restrictive Cobb-Douglas functional form in the production function approach; and it enables one to assess whether the amount of public infrastructure is insufficient or excessive by comparing the shadow value of public infrastructure and its market price. Empirical results from such models suggest smaller, but statistically significant

and more robust, estimates of infrastructure effects on overall productivity growth than found in studies employing the production function approach.

In this paper, we follow the cost function approach. Unlike previous (cost function approach) studies, however, we follow the classical papers by Meade (1952) and Hulten and Schwab (1991) and make a distinction between two channels through which public infrastructure affects economic performance. According to the first channel, as Hulten and Schwab (1991, p. 124) put it, government capital

acts as an "environmental" factor which enhances the productivity of some or all of the private inputs. Such environmental factors are essentially externalities in the sense of Romer (1986) and Lucas (1988), and seem to correspond to Aschauer's indirect effects and transportation analysts' "system effects."

More explicitly, the first channel can be expressed by specifying the Hicksneutral efficiency term, A, as a function of public infrastructure, that is, $A = A(g, \dots)$. In the more general case where the assumption of Hicks-neutrality is relaxed (which is what we do in this paper), the second channel can be formulated as $A_i = A_i(g, \dots)$, where A_i is the efficiency index for private input *i*. Apparently, in this case, public infrastructure enters the firm's production (cost) function as a determinant of total factor productivity. According to the second channel, public capital can enter the production process as a direct, but unpaid, factor of production. As discussed by Hulten and Schwab (1991), in the case of the specification of a production function, the second channel implies that public capital should be treated symmetrically with private inputs (e.g., private capital and labor).

To distinguish between these two channels, we refer to the effect of public infrastructure through the first channel as a "productivity effect" or "spillover effect" and to that through the second channel as a "production factor effect." In this paper, we focus on deriving the productivity effect of public capital. Compared with the production factor effect, the productivity effect is generally more important. This is because it is this effect that leads to divergence between marginal social benefits and marginal social costs if infrastructure is left entirely up to the market, and it is also because it is this effect that makes the study of public infrastructure interesting to policy makers and researchers.

However, we cannot disentangle the productivity effect from the production factor effect on theoretical grounds. To overcome this problem, in this paper we follow Hulten and Schwab (1991) by concentrating on an important subsector of the economy in which the effects of public capital are likely to be confined to the productivity effect [or "indirect effect" in the terminology of Hulten and Schwab (1991)]—manufacturing. As argued by Hulten and Schwab (1991, pp. 125–126),

public capital is a direct input to the transportation and communication sectors, to public utilities, and to some service industries. These sectors then pass along these direct services by selling their output to other industries. As a result, public capital does not enter the production function of other sectors as a direct input, but as a purchased intermediate good. The direct productivity of this public capital is thus accounted for through the productivity of the purchased inputs, and its value is reflected in the payments for those intermediate inputs. For sectors like

1370 GUOHUA FENG AND APOSTOLOS SERLETIS

manufacturing, any correlation between private output and public capital should in principle reflect only indirect environmental (i.e., productivity) effects.

In our particular case, we construct the price and quantity series for intermediate materials by aggregating energy (E), materials (M), and purchased business services (S), using the Fisher ideal index. Thus the production factor effect is accounted for through this aggregated input of intermediate materials (i.e., the aggregate of E, M, and S), and the effects of public capital in the manufacturing industries should in principle reflect the productivity effect. It also should be noted that in practice, the input data for manufacturing industries are not entirely purged of the production factor effect of public capital. This is because manufacturing industries provide a certain amount of transportation services (by trucks and autos) to themselves and the stocks of highways and roads should enter the manufacturing production function as part of the input used in generating the self-provided transportation services. However, given the small weight assigned to the vehicles, this impact is expected to be very small. See Hulten and Schwab (1991, pp. 133– 134) for more details. In other words, the production factor (or direct) effects of public capital are not going to bias the estimates of the productivity effects.

Motivated by the widespread practice of ignoring the distinction between the two effects of public infrastructure, the purpose of this paper is to reinvestigate the effect of public infrastructure on the performance of U.S. manufacturing industry. In doing so, we use the KLEMS data over the period from 1953 to 2001 and focus on the more interesting productivity or spillover effect—the first channel of public infrastructure. Moreover, for the first time in this literature, we employ the globally flexible AIM (asymptotically ideal model) cost function, introduced by Barnett et al. (1991) and extended by Feng and Serletis (2008) to allow for technical change. The AIM functional form has also been used in the consumer context by Serletis and Shahmoradi (2005, 2008) and Serletis and Feng (2010).

The AIM cost function is chosen for two reasons. First, the AIM cost function with technical change, extended by Feng and Serletis (2008), is particularly suitable for a situation where an economic variable enters the firm's production function both as part of intermediate inputs (i.e., the second channel) and through the efficiency term (i.e., the first channel). In particular, technical change in the AIM cost function proposed in Feng and Serletis (2008) is essentially factor-augmenting, an assumption that has been made in many studies of U.S. manufacturing industry. The assumption of factor-augmenting technical change allows public infrastructure to enter the AIM cost function through both the first and second channels.

The second reason for using the AIM cost function is its global flexibility. Most of the existing empirical literature that deals with the correlation between public infrastructure and macroeconomic performance uses the Cobb–Douglas, translog, and generalized Leontief functional forms. As is well known, however, the Cobb– Douglas function is not flexible, in that it has a constant elasticity of substitution. In the case of the generalized Leontief, Caves and Christensen (1980) have shown that it has satisfactory local properties only when technology is nearly homothetic and substitution is low. As for the translog, Guilkey et al. (1983) show that it is globally regular if and only if technology is Cobb–Douglas. That is, the translog performs well if substitution between all factors is close to unity. In contrast to these functional forms, the AIM cost function is based on a linearly homogeneous multivariate Müntz–Szatz series expansion and is globally flexible, in the sense that it is capable of approximating the underlying cost function at every point in the function's domain by increasing the order of the expansion.

The rest of the paper is organized as follows. In Section 2, we derive the general price-augmenting cost function, where public infrastructure is incorporated as part of the efficiency indexes. In Section 3, we present the price-augmenting AIM cost function, in which the efficiency indices are assumed to take the form of a Box–Cox function. Section 4 deals with the econometric specification and estimation issues. Section 5 provides a description of U.S. manufacturing data and presents and discusses the empirical results. The final section summarizes and concludes the paper.

2. THE PRICE-AUGMENTING COST FUNCTION

Before proceeding to the definition of the price-augmenting cost function, we first define its dual production function. Under the assumption of non–Hicks neutral technical change, the production function can be written as

$$y = f(X_1, X_2, \dots, X_n) = f[A_1(g, t)x_1, \dots, A_n(g, t)x_n],$$
 (1)

where y is output, f a continuous twice differentiable nondecreasing and quasiconcave function, g public infrastructure, t a time trend, x_i $(i = 1, \dots, n)$ is the *i*th actual input measured in conventional units, and X_i is the *i*th effective input. X_i is related to x_i through the functional relationship

$$X_i = A_i(g, t) x_i, i = 1, \cdots, n,$$
 (2)

where $A_i(g, t)$ is an efficiency index associated with the *i*th input. To avoid notational clusters, $A_i(g, t)$ is suppressed into A_i in what follows. Apparently, $f(A_1x_1, \dots, A_nx_n)$ in (1) is a factor-augmenting production function.

Within this primal setup, technical change (or total factor productivity growth, TFPG for short), defined as a shift in the production frontier, can easily be shown to be

TFPG =
$$\frac{\partial \ln f}{\partial t} = \sum_{i=1}^{n} \eta_i \frac{A_i}{A_i},$$

where $\eta_i = \partial \ln f / \partial \ln x_i$ is the elasticity of output with respect to input x_i and the dot above A_i denotes the change in A_i over time. However, estimation of the production function is likely to suffer from several problems, including

endogeneity, lack of flexibility, lack of economic content, and exclusion of intermediate materials when they are substitutable to private capital and labor—see, for example, Berndt and Hansson (1992), Caves et al. (1982), and Morrison and Siegel (1999). To avoid these problems, in what follows we derive a general price-augmenting cost function that is dual to the factor-augmenting production function in equation (1).

In doing so, we assume that a firm is minimizing its total cost, as follows:

$$\min_{\{x_1,\dots,x_n\}} C = \sum_{i=1}^n p_i x_i,$$
(3)

subject to

$$y = f(A_1x_1, A_2x_2, \cdots, A_nx_n),$$

where *C* is the total cost and p_i $(i = 1, \dots, n)$ represents the price for the *i*th input in conventional units. As in equation (2), the actual quantities x_i in (3) can be augmented into effective quantities X_i by the efficiency index A_i . Previous studies in the macroeconomics literature using the factor-augmenting production function approach are usually vague about the determinants of A_i , typically assuming that A_i is a function of the time trend, *t*. We depart from that approach, and in this paper we explicitly assume that A_i is a function of public infrastructure, *g*, and the time trend, *t*, the latter used as a proxy for other technology. Formally, we assume that

$$A_i = A_i(g, t) \tag{4}$$

for $i = 1, \dots, n$. Hence, in our formulation, public infrastructure, g, plays the role of increasing the efficiency of private factor inputs (i.e., capital, labor, energy, and materials).

Define $P_i \equiv p_i/A_i$, for $i = 1, \dots, n$, to be the effective price of X_i such that $p_i x_i = P_i X_i$. The inputs and their corresponding prices in conventional units can then be written as

$$x_i = \frac{X_i}{A_i},\tag{5}$$

$$p_i = A_i P_i, \tag{6}$$

respectively, for $i = 1, \dots, n$. Substituting equations (5) and (6) into (3) gives a minimization problem equivalent to that in (3):

$$\min_{\{X_1, \cdots, X_n\}} C = \sum_{i=1}^n P_i X_i,$$
(7)

subject to

 $y = f(X_1, X_2, \cdots, X_n).$

The optimal solution to the minimization problem (7), X_i^* , for $i = 1, 2, \dots, n$ is a function of prices and y; formally, $X_i^* = X_i(P_1, \dots, P_n, y)$. Substituting X_i^* into the objective function in (7) gives the following (general) efficiency index–augmented cost function, which is dual to the factor-augmenting production function in (1):

$$C = C(y, P_1, P_2, \dots, P_n) = C\left(y, \frac{p_1}{A_1}, \frac{p_2}{A_2}, \dots, \frac{p_n}{A_n}\right),$$
(8)

where the second equality is obtained by using (6), and the asterisk superscript indicating the optimal cost is dropped for simplicity. According to equation (8), the efficiency index–augmented cost function is a function of output, y, and effective input prices, p_i/A_i ($i = 1, \dots, n$). We call a cost function with effective input prices the "price-augmenting cost function," to reflect the fact that efficiency-adjusted private input prices are used.

Using the envelope theorem,

$$\frac{\partial C}{\partial t} = -\frac{\partial C}{\partial y} \frac{\partial f (A_1 x_1, A_2 x_2, \cdots, A_n x_n)}{\partial t},$$

technical change (or total factor productivity growth) can then be measured from the cost function as

$$TFPG = \frac{\partial \ln f (A_1 x_1, A_2 x_2, \dots, A_n x_n)}{\partial t}$$

$$= \frac{1}{y} \frac{\partial f (A_1 x_1, A_2 x_2, \dots, A_n x_n)}{\partial t} = -\frac{\partial C/\partial t}{y \partial C/\partial y} = -\frac{\partial \ln C/\partial t}{\partial \ln C/\partial \ln y} = -\epsilon_{ct} \epsilon_{cy}^{-1},$$
(9)

where $\epsilon_{cy} = \partial \ln C / \partial \ln y$ and $\epsilon_{ct} = \partial \ln C / \partial t$ —see also Feng and Serletis (2008) for a similar derivation of (10). Using Shephard's lemma,

$$x_i = \frac{\partial C}{\partial p_i}, \qquad i = 1, \cdots, n,$$
 (10)

we can also obtain (see Appendix A)

$$\epsilon_{ct} = \partial \ln C / \partial t = \sum_{i=1}^{n} s_i \frac{\dot{A_i}}{A_i}, \qquad (11)$$

where $s_i = p_i x_i / C$ is the cost share of input *i*.

According to equation (10), total factor productivity growth is the negative of the product of the dual rate of cost diminution, ϵ_{ct} , which is the average of the growth rates of efficiency levels weighted by their respective input cost shares, s_i , and the dual rate of returns to scale, ϵ_{cy}^{-1} . Under constant returns to scale, $\epsilon_{cy} = 1$, and total factor productivity is the negative of the dual rate of cost diminution, meaning that a 1% upward shift in the production function is equal to a 1% decrease in the cost of production. In our formulation, as part of A_i , public infrastructure, g, is a factor

that determines productivity growth, rather than being a public fixed input that is symmetric with fixed private inputs, as in previous studies.

3. MODEL SPECIFICATION

The parametric analysis of the effect of public infrastructure on private sector performance within the framework of the price-augmenting cost function in (8) requires the specification of two elements—the cost function, $C(y, p_1/A_1, \dots, p_n/A_n)$, and the efficiency index, A_i $(i = 1, \dots, n)$.

3.1. The Price-Augmenting AIM Cost Function

There are many alternatives for the functional form of the cost function in equation (8). For example, Feng and Serletis (2008) present an empirical comparison and evaluation of the effectiveness of four well-known flexible cost functions—the locally flexible generalized Leontief [see Diewert (1971)], translog [see Christensen et al. (1975)], and normalized quadratic [see Diewert and Wales (1987)], and one globally flexible cost function, the AIM [see Barnett et al. (1991)]. Another globally flexible functional form is Gallant's (1982) Fourier flexible functional cost form, based on the Fourier series expansion, recently used by Feng and Serletis (2009).

Both the Fourier and AIM globally flexible forms are capable of approximating the underlying cost function at every point in the function's domain by increasing the order of the expansion, and thus have more flexibility than most of the locally flexible functional forms, which theoretically can attain flexibility only at a single point or in an infinitesimally small region. However, as noted by Barnett and Serletis (2008), in this literature there is no a priori view as to which flexible functional forms are appropriate, once they satisfy the theoretical regularity conditions of neoclassical microeconomic theory—positiveness, monotonicity, and curvature. With this in mind, in this study we employ the AIM cost functional form, in an effort to extend our earlier work in this area [see Feng and Serletis (2008)].

One of the features of the KLEMS data set used in this paper is that constant returns to scale have been built in by the U.S. Bureau of Labor Statistics.¹ To be consistent with this feature, the (general) price-augmenting cost function in (8) has been assumed to take the form of $C(\mathbf{p}, y, g, t) = yc(\mathbf{p}, g, t)$, where $c(\cdot)$ is the unit cost function. Assuming that the unit cost function takes the form of multivariate Müntz–Szatz series expansion, we get the price-augmenting AIM total cost function

$$C = C(\mathbf{p}, y, g, t) = y \left[\sum_{z \in A_{\kappa}} b_z \prod_{j=1}^{2\kappa} \left(\frac{p_{i_j}}{A_{i_j}} \right)^{2^{-\kappa}} \right] = y \left[\sum_{z \in A_{\kappa}} b_z \prod_{j=1}^{2\kappa} (P_{i_j})^{2^{-\kappa}} \right],$$
(12)

where p > 0 is a vector of input prices (in conventional units), κ is the order of the expansion, b_z the unknown parameters, *n* the number of production factors, $B_{\kappa} = \{(i_1, i_2, \dots, i_{2^{\kappa}}) : i_1, i_2, \dots, i_{2^{\kappa}} \in \{1, 2, \dots, n; i_1 \le i_2, \le \dots \le i_{2^{\kappa}}\}, A$ is the efficiency index, and *P* is used to denote effective prices, as defined previously. It should be noted that (12) is an extension of the Barnett et al. (1991) AIM cost function without technical change, and also a generalization of the AIM cost function with technical change proposed by Feng and Serletis (2008).

Regardless of the specification of the efficiency index, A_{i_j} , in (12), our new AIM total cost function with price-augmenting technical change retains all the theoretical properties of the Barnett et al. (1991) AIM cost function without technical change. In particular, it is still globally flexible, in the sense that it is capable of approximating the underlying cost function at every point in the function's domain by increasing the order of expansion κ . Moreover, the sum of the exponents of prices in each term in (12) is still $2^{\kappa}2^{-\kappa} = 1$, thus satisfying the property of global linear homogeneity.

Although constant returns to scale are assumed in (12) to be consistent with the built-in feature of the BLS KLEMS data, (12) can easily be generalized to allow for arbitrary returns to scale. More specifically, there are two ways to generalize the price augmenting AIM cost function in (12). The first way, proposed by Barnett et al. (1991), is to add to (12) one additional parameter, ρ , as the power of the output, y. Formally, the resulting generalized cost function can be written as

$$C = C(\mathbf{p}, y, g, t) = y^{\rho} \left[\sum_{z \in A_{\kappa}} b_{z} \prod_{j=1}^{2\kappa} \left(\frac{p_{i_{j}}}{A_{i_{j}}} \right)^{2^{-\kappa}} \right] = y^{\rho} \left[\sum_{z \in A_{\kappa}} b_{z} \prod_{j=1}^{2\kappa} (P_{i_{j}})^{2^{-\kappa}} \right].$$
(13)

The second way, proposed by Thomsen (2000) in the context of the generalized Leontief cost function, is to introduce arbitrary returns to scale through the efficiency indexes (i.e., A_{i_j}) in (12). Formally, the cost function generalized in this way can be written as

$$C = C(\mathbf{p}, y, g, t) = \sum_{z \in A_{\kappa}} b_{z} \prod_{j=1}^{2\kappa} \left(\frac{p_{i_{j}}}{A_{i_{j}}}\right)^{2^{-\kappa}} = \sum_{z \in A_{\kappa}} b_{z} \prod_{j=1}^{2\kappa} (P_{i_{j}})^{2^{-\kappa}}, \quad (14)$$

where A_{i_j} is a function of g, t, and y; i.e., $A_{i_j} = A_{i_j}(t, g, y)$. It is easy to verify that the cost function generalized in the first way [i.e., (13)] is a special case of the one generalized in the second way [i.e., (14)]. However, because constant returns to scale have been built in the data, (12) will be used in what follows.

3.2. The Price-Augmenting Efficiency Index

We also need to specify a functional form for the efficiency index, A_i , $i = 1, \dots, n$. We may assume an exponential form with its argument linear in public

infrastructure, g, and a time trend, t, as follows:

$$A_i = \exp[h_i(t,g)] = \exp(\vartheta_i t + \gamma_i \ln g)$$
(15)

for $i = 1, \dots, n$, where ϑ_i is the constant growth rate of efficiency due to other technology for input *i* and γ_i is the constant elasticity of the total cost with respect to public infrastructure for input *i*. This specification is quite similar to that used in the macroeconomics growth literature, where the factor-augmenting efficiency index is commonly specified as an exponential function of the time trend; for example, $A_i = \exp(\vartheta_i t)$. In equation (15), $A_{i,0}$, the initial efficiency level of input *i*, is a constant and does not affect the calculation of elasticities and productivity growth, and thus is dropped from the efficiency index for notational simplicity.

Although simple and elegant, the specification in (15) lacks enough flexibility in modeling the effects of public infrastructure and other technology (*t*) on the efficiency levels of private factor inputs, because it is not clear whether these growth rates (ϑ_i and γ_i) should exhibit constant, logarithmic, or hyperbolic patterns over time. To allow more flexibility in modeling both the effect of public infrastructure and that of the time trend on the efficiency level of private inputs, we instead use a Box–Cox functional form for both the growth rate of the time trend and the cost elasticity with respect to public infrastructure. Formally,

$$A_{i} = \exp[h_{i}(t,g)]$$

$$= \exp\left\{\frac{\vartheta_{i}t_{0}}{\delta_{i}}\left[\left(\frac{t}{t_{0}}\right)^{\delta_{i}} - 1\right] + \frac{\gamma_{i}\ln g_{0}}{\lambda_{i}}\left[\left(\frac{\ln g}{\ln g_{0}}\right)^{\lambda_{i}} - 1\right]\right\},$$
(16)

where δ_i is the curvature parameter of the Box–Cox function for the time trend, t, and λ_i that for the log of public infrastructure, ln g. Note that when $\delta_i = 1$, $\delta_i = 0$, or $\delta_i < 0$, $\vartheta_i t_0 [(t/t_0)^{\delta_i} - 1]/\delta_i$ is a linear, log-linear, or hyperbolic, respectively, function in t. Similarly, when $\lambda_i = 1$, $\lambda_i = 0$, or $\lambda_i < 0$, $\gamma_i \ln g_0 [(\ln g/\ln g_0)^{\lambda_i} - 1]/\lambda_i$ is a linear, log-linear, or hyperbolic, respectively, function in $\ln g$. It should also be noted that $\ln g$ is scaled by its initial value (i.e., $\ln g_0$), and thus γ_i can be interpreted as the cost elasticity with respect to public infrastructure for input i at the beginning of the sample period. Similarly, ϑ_i can be interpreted as the growth rate of efficiency due to other technology for input i at the beginning of the sample period. Substituting (16) into (12) gives the price-augmenting AIM cost function used in this paper.

According to our theoretical discussion in Section 2, there must exist a production function dual to our price-augmenting AIM cost function defined by equations (12) and (16). This production function can be written as $y = F(A_1x_1, A_2x_2, \dots, A_nx_n)$, where $F(\cdot)$ is a specific functional form taken by $f(\cdot)$ in equation (1) and A_i is defined in equation (16). Under the assumption of Hicksian neutrality (that is, $A = A_i$, for $i = 1, \dots, n$), this production function

reduces to

$$y = AF(x_1, x_2, \dots, x_n)$$

$$= A_0 \exp\left\{\frac{\vartheta t_0}{\delta} \left[\left(\frac{t}{t_0}\right)^{\delta} - 1\right] + \frac{\gamma \ln g_0}{\lambda} \left[\left(\frac{\ln g}{\ln g_0}\right)^{\lambda} - 1\right]\right\} F(x_1, x_2, \dots, x_n),$$
(17)

where

$$A = A_i = \exp\left\{\frac{\vartheta t_0}{\delta}\left[\left(\frac{t}{t_0}\right)^{\delta} - 1\right] + \frac{\gamma \ln g_0}{\lambda}\left[\left(\frac{\ln g}{\ln g_0}\right)^{\lambda} - 1\right]\right\}$$

in (17) is essentially the standard Hicks neutral technical change in the Cobb– Douglas production function, and $F(x_1, x_2, \dots, x_n)$ is dual to the AIM cost function without technical change. This further confirms the validity of A_i in equation (16) as a measure of technical change.

Our price-augmenting cost function approach, based on our new priceaugmenting AIM cost function defined by equations (12) and (16), possesses a number of advantages over the simple Cobb-Douglas production function approach. First, it takes explicit account of the firm's cost optimization behavior by considering input quantities as endogenous variables, although treating input prices, which are more likely to be market-determined, as exogenous variables. As such, it is less likely to suffer from the problem of endogeneity. Second, it is globally flexible in that it is capable of approximating the underlying cost function at every point in the function's domain by increasing the order of the expansion, κ , whereas the Cobb–Douglas production function is very restrictive, in the sense that it imposes a priori the condition of constant elasticity of substitution among inputs. Third, it allows the measurement of input-specific cost elasticity with respect to public infrastructure, as well as the contribution of each input to overall cost elasticity, because an efficiency index is specified for each of the n inputs. Last but not least, the specification of the efficiency index as a Box-Cox function enables us to investigate the time pattern of the effects of public infrastructure on cost structure and productivity, which is unfortunately missed in most of the previous studies.

Our price-augmenting cost function approach provides a rich framework for investigating the effects of public infrastructure on private sector performance. This can be accomplished using three measures—the cost elasticity with respect to public infrastructure, the output elasticity with respect to public infrastructure, and the contribution of public infrastructure to total factor productivity growth. In addition, we can also obtain the social rate of return to public infrastructure, which can help answer the important policy question of whether public infrastructure is oversupplied or undersupplied. These measures are discussed in detail in what follows.

3.3. Cost Elasticity with Respect to Public Infrastructure

The spillover effect of public infrastructure on total cost is captured by the magnitude and sign of the cost elasticity with respect to public infrastructure, $\partial \ln C / \partial \ln g$, which represents the percentage change in total cost due to a 1% change in public infrastructure. In our particular case, the cost elasticity can be obtained (see Appendix B) by

$$\frac{\partial \ln C}{\partial \ln g} = -\sum_{i=1}^{n} s_i \left[\gamma_i \left(\frac{\ln g}{\ln g_0} \right)^{\lambda_i - 1} \right].$$
(18)

According to (18), the cost elasticity with respect to public infrastructure is an input cost-share weighted average of the input-specific cost elasticities (ISCE) with respect to public infrastructure, which can be written as follows:

$$ISCE_i = \gamma_i \left(\ln g / \ln g_0 \right)^{\lambda_i - 1}.$$
(19)

An advantage of (18), unlike previous studies on public infrastructure, is that we can measure input-specific cost elasticity with respect to public infrastructure, as well as the contribution of each input to overall cost elasticity, which can be written as

$$cg_{i} = \frac{s_{i}\gamma_{i} (\ln g / \ln g_{0})^{\lambda_{i}-1}}{\sum_{i=1}^{n} s_{i} \left[\gamma_{i} (\ln g / \ln g_{0})^{\lambda_{i}-1}\right]}.$$
 (20)

3.4. Output Elasticity with Respect to Public Infrastructure

Commonly used in the production function approach, the output elasticity with respect to public infrastructure, $\partial \ln f / \partial \ln g$, is another important measure for evaluating the effects of public infrastructure on private sector performance. Although it cannot be directly obtained within a cost function framework, it can be derived indirectly from the cost elasticity with respect to public infrastructure in (20) by exploiting the duality between the cost function and the production function. In fact, the implied output elasticity is the negative of the cost elasticity in (20) under the assumption of constant returns to scale.

To see this, we apply the envelope theorem to the cost minimization problem in (3) to obtain

$$\frac{\partial C}{\partial \ln g} = -\frac{\partial C}{\partial y} \frac{\partial f (A_1 x_1, \cdots, A_n x_n)}{\partial \ln g}.$$

The implied output elasticity of public infrastructure, $\partial \ln f (A_1x_1, \dots, A_nx_n) / \partial \ln g$, can then be measured from the cost function as follows:

$$\frac{\partial \ln f (A_1 x_1, \dots, A_n x_n)}{\partial \ln g} = \frac{1}{y} \frac{\partial f (A_1 x_1, \dots, A_n x_n)}{\partial \ln g}$$
$$= -\frac{\partial C/\partial \ln g}{y \partial C/\partial y}$$
$$= -\frac{\partial \ln C/\partial \ln g}{\partial \ln C/\partial \ln y}$$
$$= -\epsilon_{c \ln g} \epsilon_{cy}^{-1},$$

where $\epsilon_{c \ln g} = \partial \ln C / \partial \ln g$ and $\epsilon_{cy} = \partial \ln C / \partial \ln y$ is returns to scale. Hence, under constant returns to scale, $\epsilon_{cy} = 1$, the implied output elasticity of public infrastructure is the negative of its cost elasticity; that is, $\partial \ln f (A_1 x_1, \dots, A_n x_n) / \partial \ln g = -\epsilon_{c \ln g}$.

3.5. Total Factor Productivity Growth and Public Infrastructure

The contribution of public infrastructure to total factor productivity growth, when coupled with that of other technology (t), provides a third perspective regarding the effects of public infrastructure on private sector performance. Applying (9) to the cost function defined by (12) and (16) within a discrete time framework, we can obtain the total factor productivity growth at time t as

TFPG_t =
$$-\sum_{i=1}^{n} \left[s_i \left(\frac{A_{i,t+1}}{A_{i,t}} - 1 \right) \right].$$
 (21)

We can also decompose TFPG_t into two important components, the productivity growth due to public infrastructure, TFPG_t^g , and the productivity growth due to other technology, TFPG_t^t . For this purpose, we first decompose A_i in (16) into two components, A_i^g and A_i^t , defined as

$$A_i^g = \exp\left\{\frac{\gamma_i \ln g_0}{\lambda_i} \left[\left(\frac{\ln g}{\ln g_0}\right)^{\lambda_i} - 1 \right] \right\}$$

and

$$A_i^t = \exp\left\{\frac{\vartheta_i t_0}{\delta_i}\left[\left(\frac{t}{t_0}\right)^{\delta_i} - 1\right]\right\},$$

respectively. The productivity growth due to public infrastructure at time t, TFPG $_t^g$, is given by

$$\text{TFPG}_{t}^{g} = -\sum_{i=1}^{n} \left[s_{i} \left(\frac{A_{i,t+1}^{g}}{A_{i,t}^{g}} - 1 \right) \right], \tag{22}$$

and the productivity growth due to other technology at time t, $TFPG_t^t$, is given by

$$TFPG_{t}^{t} = -\sum_{i=1}^{n} \left(s_{i} \frac{\dot{A}_{i}^{t}}{A_{i}^{t}} \right)$$

$$= -\sum_{i=1}^{n} \left[s_{i} \left(\frac{\partial \ln C}{\partial t} \right)_{i} \right]$$

$$= -\sum_{i=1}^{n} \left\{ s_{i} \left[\gamma_{i} \left(\frac{t}{t_{0}} \right)^{\lambda_{i}-1} \right] \right\}.$$

$$(23)$$

A comparison can thus be made between TFPG_t^g and TFPG_t^r to see whether public infrastructure is a significant contributor to total factor productivity growth. It should be noted that TFPG is not equal to the sum of TFPG_t^g and TFPG_t^r . In fact, it is easy to show, as follows, that it is equal to the sum of TFPG_t^g and TFPG_t^r , plus a term representing the interaction between TFPG_g^g and TFPG_t^r :

$$TFPG = (1 + TFPGg)(1 + TFPGt) - 1$$

$$= TFPGg + TFPGt + TFPGg \times TFPGt.$$
(24)

3.6. The Social Rate of Return to Public Infrastructure

An important public policy question in this literature is whether public capital is over- or undersupplied. This question can be answered by resorting to the well-known Samuelson condition [see Samuelson (1954)], which requires that public capital (under the assumption of lump-sum taxation) be provided up to the point where the sum of marginal benefits to producers and consumers is equal to the marginal cost of providing an additional unit of public capital—see Kaizuka (1965). In calculating the marginal benefit and marginal cost, we follow the previous literature and ignore the benefits to consumers and complications resulting from the absence of lump-sum taxation. In other words, we can determine the marginal benefit and marginal cost of providing an additional unit of public capital based only on the production sector of the economy.

We assume that the government chooses the amount of public infrastructure by minimizing the present value of the costs of all the resources in the economy—see Nadiri and Mamuneas (1998). That is, the government selects the level of public infrastructure such that the sum of the industry marginal benefits equals the user cost of public capital; i.e.,

$$\sum_{h=1}^{H} m_{h,g} (\mathbf{p}, y, t, g) = p_g(r + \delta_g),$$
(25)

where $m_{h,g}(\mathbf{p}, y, t, g) = -\partial C/\partial g$ is the marginal benefit (or shadow value) of public infrastructure, which reflects the reduction in costs due to an incremental addition to the stock and can be obtained using the cost elasticity in (18); *h* indicates industry; *H* is the total number of industries (which is equal to one when only the aggregate manufacturing industry is considered); *r* is the discount factor; δ_g is the depreciation rate of public infrastructure; and p_g is the acquisition price. Because public sector capital formation is generally financed through taxation and has significant distortive effects on private sector decisions, p_g is the sum of the direct burden of the taxes needed to pay for the infrastructure and the deadweight cost associated with these taxes for the last dollar of public investment. Solving equation (25) for g^* yields the optimal amount of public infrastructure.

Let MB = $\sum_{h=1}^{H} m_{h,g}(\mathbf{p}, y, t, g)$ denote the marginal benefit and MC = $p_g(r + \delta)$ denote the marginal cost. Then the Samuelson condition, together with (25), implies that

Public infrastructure is
$$\begin{cases} \text{optimally supplied} & MB = MC \\ \text{undersupplied} & \text{if } MB > MC \\ \text{oversupplied} & MB < MC. \end{cases}$$
 (26)

The Samuelson condition can be also stated in terms of the net social rate of return to public infrastructure. To see this, we can rearrange (25) to obtain

$$\frac{\sum_{h=1}^{H} m_{h,g} \left(\boldsymbol{p}, y, t, g \right)}{p_g} - \delta_g = r.$$
 (27)

Letting γ_s denote the left-hand side of equation (27) (the net social rate of return to public infrastructure), equation (26), the Samuelson condition, can alternatively be written as

Public infrastructure is
$$\begin{cases} \text{optimally supplied} & \gamma_s = r \\ \text{undersupplied} & \text{if } \gamma_s > r \\ \text{oversupplied} & \gamma_s < r, \end{cases}$$
 (28)

where r and δ_g are defined as before. Equation (28) is the Samuelson condition we will use in this paper in answering the question of whether public capital is over- or undersupplied. However, as can be seen from (27), the application of (28) requires the specification of p_g , δ_g , and r, which will be discussed in Section 5.6.

4. ECONOMETRIC SPECIFICATION AND ESTIMATION ISSUES

In empirical applications, the approximation of the AIM cost function must be truncated at some finite value κ (i.e., finite partial sums). The order of approximation κ is usually determined empirically and stops when the elasticity estimates and the covariance matrix of the disturbances converge. In this paper, because of degree-of-freedom problems, we set $\kappa = 2$. Hence, with n = 3 (the case in this paper) and $\kappa = 2$ in equation (12), we get the price-augmenting AIM(2) cost function

$$C_{\kappa=2} (\mathbf{p}, \mathbf{y}, \mathbf{g}, t) = \mathbf{y} \Big(b_1 P_1 + b_2 P_2 + b_3 P_3 + b_4 P_1^{1/2} P_2^{1/2} + b_5 P_1^{1/2} P_3^{1/2} + b_6 P_2^{1/2} P_3^{1/2} + b_7 P_1^{3/4} P_2^{1/4} + b_8 P_1^{1/4} P_2^{3/4} + b_9 P_1^{3/4} P_3^{1/4} + b_{10} P_1^{1/4} P_3^{3/4} + b_{11} P_2^{3/4} P_3^{1/4} + b_{12} P_2^{1/4} P_3^{3/4} + b_{13} P_1^{1/2} P_2^{1/4} P_3^{1/4} + b_{14} P_1^{1/4} P_2^{1/2} P_3^{1/4} + b_{15} P_1^{1/4} P_2^{1/4} P_3^{1/2} \Big).$$
(29)

Applying (10) to (29) yields the following system of factor demand equations for the AIM(2) model with n = 3:

$$\begin{aligned} \frac{x_1}{y} &= \frac{1}{A_1} \left(b_1 + \frac{1}{2} b_4 P_1^{-1/2} P_2^{1/2} + \frac{1}{2} b_5 P_1^{-1/2} P_3^{1/2} \right. \tag{30} \\ &+ \frac{3}{4} b_7 P_1^{-1/4} P_2^{1/4} + \frac{1}{4} b_8 P_1^{-3/4} P_2^{3/4} + \frac{3}{4} b_9 P_1^{-1/4} P_3^{1/4} \\ &+ \frac{1}{4} b_{10} P_1^{-3/4} P_3^{3/4} + \frac{1}{2} b_{13} P_1^{-1/2} P_2^{1/4} P_3^{1/4} + \frac{1}{4} b_{14} P_1^{-3/4} P_2^{1/2} P_3^{1/4} \\ &+ \frac{1}{4} b_{15} P_1^{-3/4} P_2^{1/4} P_3^{1/2} \right); \end{aligned}$$

$$\frac{x_2}{y} = \frac{1}{A_2} \left(b_2 + \frac{1}{2} b_4 P_1^{1/2} P_2^{-1/2} + \frac{1}{2} b_6 P_2^{-1/2} P_3^{1/2} \right)$$

$$+ \frac{1}{4} b_7 P_1^{3/4} P_2^{-3/4} + \frac{3}{4} b_8 P_1^{1/4} P_2^{-1/4} + \frac{3}{4} b_{11} P_2^{-1/4} P_3^{1/4}$$

$$+ \frac{1}{4} b_{12} P_2^{-3/4} P_3^{3/4} + \frac{1}{4} b_{13} P_1^{1/2} P_2^{-3/4} P_3^{1/4} + \frac{1}{2} b_{14} P_1^{1/4} P_2^{-1/2} P_3^{1/4}$$

$$+ \frac{1}{4} b_{15} P_1^{1/4} P_2^{-3/4} P_3^{1/2} ;$$
(31)

$$\frac{x_3}{y} = \frac{1}{A3} \left(b_3 + \frac{1}{2} b_5 P_1^{1/2} P_3^{-1/2} + \frac{1}{2} b_6 P_2^{1/2} P_3^{-1/2} + \frac{1}{4} b_9 P_1^{3/4} P_3^{-3/4} + \frac{3}{4} b_{10} P_1^{1/4} P_3^{-1/4} + \frac{1}{4} b_{11} P_2^{3/4} P_3^{-3/4} + \frac{3}{4} b_{12} P_2^{1/4} P_3^{-1/4} + \frac{1}{4} b_{13} P_1^{1/2} P_2^{1/4} P_3^{-3/4} + \frac{1}{4} b_{14} P_1^{1/4} P_2^{1/2} P_3^{-3/4} + \frac{1}{2} b_{15} P_1^{1/4} P_2^{1/4} P_3^{-1/2} \right).$$
(32)

Concavity (in prices) requires that the Hessian matrix of the second derivatives of the cost function with respect to prices, $\nabla_{p_i p_j} C(\mathbf{p}, y, g, t)$, be negative semidefinite. In practice, concavity of the cost function may not be satisfied. In that case, we impose concavity fully (that is, at every data point in the sample) on the AIM model, using methods suggested by Gallant and Golub (1984), to which we now turn.

4.1. Seminonparametric Estimation

The AIM(2) factor demand system, equations (30)–(32), can be written as

$$z_t = \psi(\mathbf{p}, y, g, t, \theta) + \epsilon_t, \qquad (33)$$

where $z = (z_1, \dots, z_n)'$ is the vector of input–output ratios, $\psi(\mathbf{p}, y, g, t, \theta)$ is given by the the right-hand sides of equations (30)–(32), and $\theta = (b_1, b_2, \dots, b_{n^{2^{\kappa}}}, \vartheta_1, \vartheta_2, \vartheta_3, \gamma_1, \gamma_2, \gamma_3, \lambda_1, \lambda_2, \lambda_3, \delta_1, \delta_2, \delta_3)$. ϵ_t is a vector of stochastic errors and we assume that $\epsilon \sim N(\mathbf{0}, \Omega)$, where **0** is a null matrix and Ω is the $n \times n$ symmetric positive definite error covariance matrix. The same assumption about the error term, ϵ , has also been made by Berndt and Hansson (1992), Nadiri and Mamuneas (1994), and Morrison and Schwartz (1996) in the public infrastructure and productivity literature, and by Diewert and Fox (2008), among many others, in the broader literature of demand systems.

As Gallant and Golub (1984, p. 298) put it,

all statistical estimation procedures that are commonly used in econometric research can be formulated as an optimization problem of the following type [Burguete, Gallant, and Souza (1982)]

$$\boldsymbol{\theta}$$
 minimizes $\varphi(\boldsymbol{\theta})$ over Θ (34)

with $\varphi(\theta)$ twice continuously differentiable in θ .

Notice that $\psi(\mathbf{p}, y, g, t, \theta)$ is nonlinear in $\vartheta_1, \vartheta_2, \vartheta_3, \gamma_1, \gamma_2$, and γ_3 , and therefore the AIM(2) factor demand system in (33) can be fitted using Gallant's (1975) seemingly unrelated nonlinear regression method to estimate θ . Hence, $\varphi(\theta)$ has the form

$$\varphi(\theta) = \frac{1}{T} \epsilon'_t \epsilon_t$$

$$= \frac{1}{T} \sum_{t=1}^T (z_t - \psi(\cdot))' \widehat{\Omega}^{-1} (z_t - \psi(\cdot)),$$
(35)

where $\widehat{\Omega}$ is an estimate of the error variance–covariance matrix, Ω . In minimizing (35), we use the TOMLAB/NPSOL tool box with MATLAB. NPSOL uses a sequential quadratic programming algorithm and is suitable for both unconstrained and constrained optimization of smooth (that is, at least twice continuously differentiable) nonlinear functions.

4.2. Endogeneity

One issue concerning our stochastic specification is the possible endogeneity of input prices and public infrastructure. To address this potential problem, in this paper we first use the Zellner method of estimation, assuming that both input prices and public infrastructure are exogenous, and then use a two-stage least squares approach (in Section 5.7) to test the robustness of our results obtained from the Zellner method.

4.3. Econometric Regularity

Another issue is that of nonstationarity. If the errors are nonstationary, then there is no theory linking the left-hand side to the right-hand side variables in equation (33) or, equivalently, no evidence for the theoretical models in level form. In such cases, some important nonstationary variables might have been omitted. Allowing for first-order serial correlation, as is usually done in the literature, is almost the same as taking first differences of the data if the autocorrelation coefficient is close to unity. In that case, the equation errors become stationary, but there is no theory for the models in first differences. Munnell (1992), in a discussion on infrastructure investment and economic growth, has argued that first differencing destroys the long-term relationships in the data and therefore it does not make economic sense to use equations in this form. In fact, previous studies in this literature that have estimated equations in first differences have found private capital and labor to be nonsignificant-see, for example, Hulten and Schwab (1991) and Sturm and de Haan (1995). Although the contribution of infrastructure can be questionable, the role played by labor and private capital is not in doubt. Duggal et al. (1999) have argued that the fact that first-differenced equations generate nonsignificant estimates for the labor and capital coefficients is enough reason to question the validity of using first differences of the data.

4.4. Theoretical Regularity

Finally, in the estimation of (33), we pay special attention to the theoretical regularity conditions of positiveness, monotonicity, and curvature. The regularity conditions are checked as in Feng and Serletis (2008), as follows:

• Positiveness is checked by checking if the estimated cost is positive,

$$C\left(\boldsymbol{p}, y, g, t\right) > 0.$$

- Monotonicity is checked by direct computation of the values of the first gradient vector of the estimated cost function with respect to p. It is satisfied if $\nabla_p C(p, y, g, t) > 0$.
- Curvature requires the Hessian matrix of the cost function to be negative semidefinite and is checked by performing a Cholesky factorization of that matrix and checking whether the Cholesky values are nonpositive [because a matrix is negative semidefinite if its Cholesky factors are nonpositive—see Lau (1978, Theorem 3.2)]. Curvature can also be checked by examining the eigenvalues of the Hessian matrix, provided that the monotonicity condition holds. It requires that these eigenvalues be negative or zero.

We first run an unconstrained optimization using (34). If theoretical regularity is not attained, then we impose the theoretical regularity conditions.

5. DATA AND EMPIRICAL EVIDENCE

5.1. Data Description

We use annual data for capital, labor, and intermediate materials for total manufacturing industry in the United States over the period from 1953 to 2001. We also use annual data for capital, labor, and intermediate materials for each of the 12 two-digit manufacturing industries to check the robustness of our results. The 12 two-digit manufacturing industries chosen are exactly the same as in Nadiri and Mamuneas (1994) and are listed in Table 1. It is to be noted that we constructed the price and quantity series for intermediate materials by aggregating energy, materials, and purchased business services, using the Fisher ideal index. All data on quantities and prices were obtained from the Bureau of Labor Statistics (BLS) at www.bls.gov/data/home.htm. We normalized all the price series to be equal to 1 in 1953 and obtained the quantity series for each of output, capital, labor, energy, materials, and purchased business services by dividing the value of production or factor cost by the corresponding normalized price series. This BLS (KLEMS) data set has been used previously by Nadiri and Mamuneas (1994), Diewert and Fox (2008), and Feng and Serletis (2008), among many others.

A major feature of the BLS data set is that constant returns to scale are built in by constructing input factor payments in such a way that they add up to the value of output. Thus, tests of returns to scale and scale bias are inappropriate, as are

SIC	Industry		
20	Food		
26	Paper and allied products		
28	Chemicals and allied products		
29	Petroleum refining and related industries		
30	Rubber products		
32	Stone, clay, and glass products		
33	Primary metals		
34	Fabricated metal products		
35	Machinery		
36	Electrical equipment		
37	Transportation		
38	Scientific instruments		

 TABLE 1. SIC classification

some tests of imperfect competition. Another feature of the BLS data set is that it provides the price and quantity series for purchased business services inputs. Directly collected data on purchased business services are relatively scant, and for that reason they have been ignored by similar studies in the past. However, there is ample evidence of an increased use of purchased business services by industries over the postwar period, and there are two important issues to consider. The first is that a sizable and growing input should not be ignored in productivity measurement, if aggregate inputs are not to be underestimated and mismeasured. The other is the possibility of substitution between capital, labor, and services purchased from outside. Examples of the latter are the substitution of leased equipment for owned capital and purchased accounting for services performed by payroll employees.

As in Duggal et al. (1999, 2007), we restrict our analysis to core public infrastructure, which consists of the following three categories: (i) highways and streets; (ii) other buildings (which include police, fire stations, court houses, auditoriums, and passenger terminals); and (iii) other structures (which include electric and gas facilities, transit systems, and airfields). *g* represents the net capital stock (net of depreciation) held by the federal, state, and local governments, expressed in billions of 1953 dollars. The data on the stock of public infrastructure are obtained from the Bureau of Economic Analysis (at http://www.bea.gov/national/FA2004/SelectTable.asp).

5.2. Theoretical Regularity Tests

In the first column of Table 2, we present the parameter estimates and theoretical regularity violations for the unconstrained AIM(2) system for U.S. total manufacturing industry. As results in nonlinear optimization are sensitive to the initial parameter values, to achieve global convergence, we randomly generated 500 sets

	Unconstrained	Curvature constrained	
Parameter		In 1977	Fully
γ_1	0.4462	0.7820	0.4396
γ_2	0.0000	0.0000	0.1683
γ3	0.7391	0.1552	0.0720
ϑ_1	0.0000	0.0045	0.0001
ϑ_2	0.0296	0.0006	0.0000
ϑ_3	0.0207	0.0240	0.0353
λ_1	5.0257	2.4428	-0.4220
λ_2	0.9904	-1.9319	-2.9598
λ_3	-4.1800	-0.2179	-6.0000
δ_1	0.1066	0.3954	1.5388
δ_2	-1.8550	-0.9063	0.4509
δ_3	1.0060	0.9694	0.8723
b_1	1.2792	1.3615	3.4984
b_2	2.4103	2.2261	2.8481
b_3	1.7761	3.8369	2.3374
b_4	4.4890	0.5454	5.1981
b_5	-6.1288	4.9277	2.9666
b_6	18.0862	17.7346	19.4638
b_7	-3.5560	-2.0610	-4.4887
b_8	8.5529	0.5853	4.9279
b_9	-4.4298	-0.1552	-6.1954
b_{10}	-12.7338	-12.5225	-11.5625
b_{11}	1.9394	-3.9858	-5.2340
b_{12}	-9.7623	-10.2151	-11.0440
b_{13}	2.2237	-0.5518	5.9160
b_{14}	4.9940	7.0677	7.0922
b_{15}	-8.1198	-7.7834	-14.7037
$\varphi(\widehat{oldsymbol{ heta}})$	0.0442	0.0448	0.0459
Positivity violations	0	0	0
Monotonicity violations	0	0	0
Curvature violations	12	4	0

TABLE 2. Price-augmenting AIM(2) parameter estimates

Note: Sample period, annual data 1953-2001 (T = 49).

of initial parameter values and chose the starting θ that led to the lowest value of the objective function. It is also to be noted that a parametric bootstrapping method is usually used in constrained optimization to obtain statistical inference for the estimated parameters ($\hat{\theta}$) or nonlinear transformations of these parameters ($\phi(\hat{\theta})$, i.e., elasticities)—see Gallant and Golub (1984). This involves the use of Monte Carlo methods, generating a sample from the distribution of the inequality-constrained estimator ($\hat{\theta}$) large enough to provide a reliable estimate of the sampling distributions of ($\hat{\theta}$) and $\phi(\hat{\theta})$. However, for computational reasons, this is unaffordable at present. Therefore, only point estimates are provided for the estimated parameters $(\hat{\theta})$ in Table 2.

As can be seen in the first column of Table 2, although positiveness and monotonicity are satisfied for all sample observations, curvature is violated for 12 data points. Because regularity has not been attained, we follow Feng and Serletis (2008, 2009) and use the NPSOL nonlinear programming program to minimize $\varphi(\theta)$, subject to the constraint that the three eigenvalues of the Hessian matrix, H, are nonpositive. This is because a necessary and sufficient condition for the concavity of H is that all its eigenvalues are nonpositive—see, for example, Morey (1986). Thus, our constrained optimization problem is written as

$$\min_{\boldsymbol{\theta}} \varphi(\boldsymbol{\theta}),$$

subject to

 $\varphi_i(\mathbf{p}, y, g, t, \boldsymbol{\theta}) < 0, \text{ for } i = 1, \cdots, n,$

where $\varphi_i(\mathbf{p}, y, g, t, \theta)$, $i = 1, \dots, n$, are the eigenvalues of the Hessian matrix of the AIM(2) cost function.

With the constrained optimization method, we can impose curvature restrictions at any arbitrary set of points—at a single data point, over a region of data points, or fully (at every data point in the sample). We minimize $\varphi(\theta)$ subject to the constraint that the cost function is locally concave at 1977 and also subject to the constraint that it is fully concave (concave at every data point). The results are reported in the second and third columns of Table 2—the second column shows the results when the curvature constraint is imposed locally (at 1977) and the third column shows the results when the constraint is imposed at every data point in the sample. Clearly, the effect of imposing the curvature constraint locally is unsatisfactory, as the number of curvature violations drops from 12 to 4. However, as we expect, the imposition of the curvature violations to zero, producing parameter estimates that are consistent with all three theoretical regularity conditions at every data point in the sample, that is, fully. Thus, in what follows we focus on results from the AIM(2) cost function with the curvature conditions imposed fully.

As seen from the third column of Table 2, the estimates of γ_i (i = 1, 2, 3) are all positive (i.e., $\gamma_1 = 0.4396$, $\gamma_2 = 0.1683$, and $\gamma_3 = 0.0720$). With ln g and ln g_0 also being positive (because both g and g_0 are greater than one), this implies that, regardless of the signs of the λ_i 's, the effect of public infrastructure on the efficiency of private inputs is always positive, in the sense that it is capable of reducing total cost [see (18)]. However, the negative signs of the curvature parameters for public infrastructure (i.e., $\lambda_1 = -0.4220$, $\lambda_2 = -2.9598$, and $\lambda_3 = -6.0000$) imply that the ability of public infrastructure to increase the efficiency level of private inputs has declined over the sample period, which is reflected hereafter in the change in the time patterns of the estimates of cost elasticities, output elasticities, productivity growth, and the social rate of return to public infrastructure.



FIGURE 1. Cost elasticity with respect to public infrastructure in total U.S. manufacturing industry.

5.3. Cost Elasticities

An important question in the literature is whether public infrastructure is productive. This question can be answered by looking at the effect of public infrastructure on cost reduction. Using equation (18), we calculated the estimates of the cost elasticity with respect to public infrastructure for total U.S. manufacturing industry, shown in Figure 1. Apparently, the estimates are negative over the sample period, with an average of -0.0993, suggesting that public infrastructure is productive in terms of cost reduction. Our estimates of the cost elasticities are roughly consistent with those made by Nadiri and Mamuneas (1994). In particular, they applied a translog cost function, treating public infrastructure as a public fixed input, to the data on the same 12 two-digit manufacturing industries (over the period from 1970 to 1986) and found that the mean values of the estimates of the cost elasticity range between -0.1000 and -0.1500 across the two-digit manufacturing industries. As we noted in Section 2, the cost reduction is actually achieved through the ability of public infrastructure to increase the efficiency levels of private inputs (capital, labor, and materials).

We are also interested in the time pattern exhibited by our estimated cost elasticity. It is clear from Figure 1 that there is a break in the estimated cost elasticity in 1973. In particular, over the period from 1953 to 1973, the estimates of cost elasticity are less negative over time, and decrease rapidly from 0.1760 in 1953 to 0.0855 in 1973 in absolute terms, meaning that a 1% increase in



FIGURE 2. Cost elasticities with respect to public infrastructure for two-digit manufacturing industries in the United States.

public infrastructure results in a smaller percentage reduction in cost over time. In the second subperiod, from 1974 to 2001, the estimates of the cost elasticity are relatively stable, with an average of -0.0750, suggesting that a 1% increase in public infrastructure leads to a 0.0750% cost reduction on average. The time pattern exhibited by the estimated cost elasticity implies that although public infrastructure offers significant benefits in terms of cost reduction in the pre-1973 period, it cannot offer the same benefits at the margin.

To further test the robustness of the result regarding the time pattern exhibited by our estimated cost elasticity in the total manufacturing industry, we also estimate the model (equation (33)) for each of the 12 two-digit manufacturing industries separately, and calculate the cost elasticity for each of them. As can be seen clearly from Figure 2, a similar time pattern (i.e., declines prior to 1974 and relative stability after that) is also found in almost all of the 12 two-digit manufacturing industries (the only exception is ISC 33). The estimates from the 12 two-digit manufacturing industries further confirm the conclusion that the externality effect of public infrastructure on cost reduction in U.S. manufacturing is more significant prior to 1974 and less important after that. In addition, we note that the estimates of the cost elasticity vary across industries, but for a majority of the industries the estimates range between -0.0600 and -0.3065 in 1953 and between -0.0100 and -0.1255 in 2001. We also note from Figure 2 that ISC 37 (Transportation Equipment), which is an intensive user of the core public infrastructure, experienced a larger decline than any of the other industries. In addition, the averages of the cost elasticities with respect to public infrastructure



FIGURE 3. Input-specific elasticities with respect to public infrastructure in total U.S. manufacturing industry.

for the 12 two-digit manufacturing industries ranges from -0.0232 to -0.1341, slightly lower than those found by Nadiri and Mamuneas (1994).

Unlike previous studies, our price-augmenting technical change approach enables us to further investigate the input specific cost elasticities with respect to public infrastructure, $\gamma_i (\ln g / \ln g_0)^{\lambda_i - 1}$, for $i = 1, \dots, n$. Two results emerge from Figure 3. First, public infrastructure is mainly capital- and labor-saving. In particular, the capital-specific cost elasticity is very impressive, averaging -0.328per year over the sample period. For labor and materials, the input-specific cost elasticities are found to be moderate, averaging -0.079 and -0.021, respectively. These results imply that cost savings are realized through the ability of public infrastructure to increase the efficiency of private inputs (capital, labor, and materials).

We further calculate the contribution of each factor to the total cost elasticity with respect to public infrastructure, using equation (18). It is clear from Figure 4 that capital and labor have been the two dominant factors causing the externality effect of public infrastructure on cost reduction. Materials have a positive and small impact on the externality effect of public infrastructure. Second, like the estimates of the total cost elasticities, all three input-specific cost elasticities exhibit a similar time pattern; i.e., they all become less negative over time in the first subperiod (the pre-1973 period) and become relatively stable after 1974. In particular, the capital-specific cost elasticity decreases from 0.440 to 0.266 in absolute terms; the labor-specific cost elasticity decreases from 0.168 to 0.041 in absolute terms; and the materials-specific cost elasticity decreases from 0.072 to 0.006 in absolute terms. The time pattern exhibited by the estimates of input-specific cost elasticities.



FIGURE 4. Contribution of input factors to total cost elasticity in total U.S. manufacturing industry.

suggests that the ability of public infrastructure to increase the efficiency of private inputs has greatly diminished over time, and that at the end of the sample period, public infrastructure is mainly capital-saving.

5.4. Output Elasticities

A further question in this literature is how productive public infrastructure is. This question can be answered by making a comparison between public infrastructure and the three private inputs in terms of output elasticity with respect to public infrastructure. As discussed in Section 3, the implied output elasticity is the negative of the cost elasticity in equation (18) under the assumption of constant returns to scale. Further, under the same assumption, it is easy to show that the output elasticities of private capital, labor, and materials can be approximated by their corresponding revenue shares.

Figure 5 plots the output elasticity of public infrastructure and also the output elasticities of private capital, labor, and materials in the total U.S. manufacturing industry. Roughly speaking, over the period 1953 to 1974, the output elasticity with respect to labor is the largest, followed by those of materials, private capital, and public infrastructure. Over the sample period 1974 to 2001, the output elasticity with respect to materials is the largest, followed by those with respect to labor, private capital, and public infrastructure. Though the smallest in both periods, the average output elasticity of public infrastructure (0.0993) is by no means small,



FIGURE 5. Output Elasticities in total U.S. manufacturing industry.

compared to those of capital, labor, and materials (0.1750, 0.4290, and 0.3962, respectively) over the sample period.

It is more instructive to compare the output elasticities of two types of capital, public infrastructure and private capital. It is clear from Figure 5 that, although the output elasticity of private capital is relatively stable over time, that of public infrastructure has declined steadily over the sample period. In particular, the output elasticity of public infrastructure is 0.1760, slightly higher than that of private capital (0.1615) at the beginning year of the sample period (i.e., in 1953 and 1954). However, it declines steadily and stabilizes at around 0.0750 after 1974. At the end of the sample period (in 2001), the output elasticity of public infrastructure is 0.0635, which is only 40% that of private capital. In other words, public infrastructure is getting less productive relative to private physical capital over the sample period.

5.5. Productivity Decomposition

A very important policy question in this literature is whether expanded investment in public infrastructure can lead to sustainable productivity growth in the U.S. economy. This question can be answered by assessing the relative importance of public infrastructure to total factor productivity growth over time. As noted in equation (24), total factor productivity growth, TFPG, can be decomposed into three components—productivity growth due to public infrastructure, TFPG^{*g*}, productivity growth due to other technology, TFPG^{*t*}, and an interaction term between public infrastructure and other technology. TFPG, TFPG^{*g*}, and TFPG^{*t*} are calculated using equations (21), (22), and (23), respectively. Because of the



FIGURE 6. Productivity growth decomposition in total U.S. manufacturing industry.

negligible magnitude of the third component, $TFPG^t$, our analysis focuses on the first two components.

Before making a comparison between the contribution of public infrastructure to TFPG and that of other technology, we first examine the performance of our model in terms of its flexibility in modeling TFPG (the solid line in Figure 6). As is well known, the commonly used general time trend approach, where the time trend is treated symmetrically with other inputs within the flexible functional forms framework, does not have enough flexibility in capturing the ups and downs in productivity growth. For example, Kohli (1990) and Diewert and Wales (1992) have noted that the general time trend approach to obtaining productivity growth estimates yields values that are smoothed versions of the estimates from the commonly used index number approaches. Moreover, Feng and Serletis (2008) have noted that even the AIM cost function with a linear time trend introduced in the efficiency index does not perform well enough in capturing the ups and downs in productivity growth; i.e., it produces estimates that can only be regarded as smoothed versions of that from the Fisher ideal index. Compared with the previous models, our model performs much better in that it captures most of the productivity ups and downs such as the productivity slowdown in the early 1970s. The exceptional performance of our model is also reflected in its flexibility in capturing the ups and downs of $TFPG^{g}$ (see the dotted line in Figure 6) and $TFPG^{t}$ (see the dashed line in Figure 6).

We now turn to the comparison between the contribution of public infrastructure to TFPG and that of other technology. It is clear from Figure 6 that TFPG^{*g*} and TFPG^{*t*} have different time patterns and magnitudes. For the productivity growth due to public infrastructure, there is a clear structural break in 1974. The TFPG^{*g*} estimates over the pre-1974 period are much greater on average than those over the post-1974 period. In particular, the estimates of TFPG^{*g*} average 0.79% over the period 1953 to 1974 whereas they average only 0.17% over the period 1974 to 2001. In fact, the decline in productivity growth due to public infrastructure partially accounts for the widely noted productivity slowdown in the early 1970s. In contrast, there is no clear structural break in 1974 for the TFPG^{*t*} estimates. Moreover, the estimates of TFPG^{*t*} over the pre-1974 period are slightly lower on average than those over the post-1974 period. In particular, the estimates of TFPG^{*t*} average 1.25% over 1953–1974 and 1.50% over 1974–2001. It is clear from Figure 6 that the productivity growth due to other technology, TFPG^{*t*}, explains most of the productivity resurgence in the late 1990s.

To see the temporal pattern of the importance of public infrastructure to productivity growth more directly, we calculate the contribution of TFPG^g to TFPG using TFPG^g/TFPG \times 100. Public infrastructure makes an impressive contribution of 33.4% in the pre-1974 period. However, its contribution declines significantly to 11.2% in the second subperiod, and to merely 4.4% in the most recent period from 1996 to 2001. In contrast, other technology contributes 66.6% to the total factor productivity growth in the pre-1974 period. Its contribution increases sharply to 88.8% in the post-1974 period, and to as high as 95.6% in the most recent period, from 1996 to 2001. Apparently, public infrastructure is a significant contributor to productivity growth in the pre-1974 period, but only a limited contributor in the post-1974 period. In other words, the manufacturing industry data do not support the view that returning public infrastructure growth to pre-1974 levels would raise productivity growth to pre-1974 levels. Taking into account the fact that the core public infrastructure is composed mainly of highways (around 60%), our finding regarding the contribution of TFPG^g to TFPG is consistent with that in Fernald (1999) that the massive infrastructure building of the 1950s, 1960s, and early 1970s-which largely reflected the construction of the interstate highway network-offered a one-time increase in the level of productivity, rather than a continuing path to productivity growth.

It should be noted that TFPG^{*g*} represents change, not level of productivity (or efficiency) due to public infrastructure. Thus, the negative value of TFPG^{*g*} in certain years (for example, 1957, 1975, 1982, 1983, 1992, 2000, and 2001) should be interpreted as a decline in the level of productivity due to public infrastructure from the previous year, rather than as a negative level of productivity due to public infrastructure. In fact, if the productivity level in 1953 is normalized to one and only productivity growth due to public infrastructure is taken into consideration, then the productivity level will be 1.0353, 1.1595, 1.1640, 1.1625, 1.1851, 1.2169, and 1.2102 for 1957, 1975, 1982, 1983, 1992, 2000, and 2001, respectively. In other words, public infrastructure has consistently

been efficiency-improving over the sample period, despite the sign changes of TFPG^{*g*}.

5.6. Rates of Return

Another important public policy question in the literature is whether public capital is over- or undersupplied. As mentioned in Section 3.6, the optimal provision of public capital services requires that the level of public capital provided be at the point where the net social rate of return to public infrastructure is equal to the discount rate—see equation (28). If the social rate of return to public infrastructure is higher (lower) than the discount rate, public infrastructure is undersupplied (oversupplied) and an increase (reduction) of public investment is necessary.

A difficulty in calculating the social rate of return in equation (28) is finding an appropriate measure of the marginal cost, p_g , because there is no consensus on the magnitude of p_g . For example, Ballard (1990) put the marginal cost of public sector investment at about \$1.20 for each dollar of benefits, whereas Jorgenson and Yun (1990) estimate this cost to be about \$1.47 for each dollar of benefits. Despite the large difference in the estimates of marginal cost of public sector investment, Browning (1987) points out that the preferred range is between 1.318 and 1.469. In this paper, we use 1.407, the median of this preferred range. As noted in Section 3.6, we also need to find appropriate measures of the depreciation rate of public infrastructure, δ_{g} , and the discount rate, r. For δ_{g} , we use the average of the depreciation rates of government nonresidential structures, weighted by the value shares in each year. The depreciation rates of government nonresidential structures can be obtained from the Bureau of Economic Analysis (at http://www.bea.gov/national/FA2004/Tablecandtext.pdf). For the discount rate, r, we use the net rate of return to private capital stock, which is approximated by that for the nonfinancial sectors in the United States-see Bureau of Economic Analysis (1999).

Figure 7 presents the net rate of return to public infrastructure and the net rate of return to private capital stock. Clearly, over the period 1953 to 1973, the rate of return to public infrastructure is much higher than that to private capital, reflecting the shortage of public infrastructure in that period. In particular, the average rate of return to public infrastructure over the period 1953 to 1973 is 20% compared with a return rate of 9.9% for private capital over the same period. However, the return to public infrastructure capital declines quickly over this period and converges to that to private capital stock in 1974. Over the period 1974 to 2001, especially after 1986, the rate of return to public infrastructure is roughly the same as that to private capital. This suggests that the rate of return to public infrastructure is are considered. The temporal pattern exhibited by the total manufacturing industries are considered. The temporal pattern exhibited by the total manufacturing industries, as shown in Figure 8.



FIGURE 7. Social rate of return to public infrastructure and rate of return to private capital in total U.S. manufacturing industry.

However, we cannot jump to the conclusion that public infrastructure was provided at its optimal level over the period 1974 to 2001. This is because the consumption sector and nonmanufacturing industries are ignored in our analysis. Moreover, the market-mediated effect of public infrastructure is not taken into



FIGURE 8. Social rate of return to public infrastructure for 12 two-digit U.S. manufacturing industries.

consideration, as discussed in the Introduction. When all these are taken into account, (27) implies that the new net rate of return to public infrastructure, r_g^w , is most likely to be greater than r_g , the net rate of return to public infrastructure when the consumption sector and nonmanufacturing industries are not considered. Because r_g is roughly the same as the rate of return to private capital in the post-1973 period, a higher rate of return to public infrastructure implies that public infrastructure is undersupplied when the whole economy is considered.

This conclusion is further justified by the effects of the Tax Reform Act of 1986 on the marginal cost of public funding, p_g . Previous studies generally find that the Tax Reform Act of 1986 reduces the marginal excess burden. For example, Jorgenson (1996, p. xxviii)) finds that the Tax Reform Act of 1986 reduces the marginal excess burden of taxation by 8%, implying that p_g should be lower by 8% in the post-1986 period than in the pre-1986 period. More specifically, if we assume that $p_g = 1.407$ for the pre-1986 period as before, then p_g in the post-1986 period should be adjusted to 1.327, implying that the net rate of return to public infrastructure in the post-1986 period is actually higher than that shown in Figure 7. Thus, taking into account both the effect of the Tax Reform Act of 1986 on p_g and the spillover effects of public infrastructure in the consumption sector and nonmanufacturing industries, public infrastructure is undersupplied even in the second subsample period.

Considering the importance of this public policy question and the lack of consensus in the economic literature with respect to the magnitude of the marginal cost of public funding [see Jacobs (2009)], in what follows we examine the sensitivity of our results regarding the rate of return to public infrastructure, using the values for p_g found by more recent studies. A central finding of the modern labor market literature is that labor supply responses tend to be concentrated along the extensive margin (labor force participation) rather than the intensive margin (hours of work), implying that p_g becomes a function of average taxes, rather than just marginal taxes. Noting this, Kleven and Kreiner (2003) recalculate p_g for 23 OECD countries, and find that it ranges from below 1 to above 2 across the sample countries. For the case of the United States, they find that p_g ranges from 0.89 to 1.34. In what follows we will examine the sensitivity of our results, using these new values for p_g .

Figure 9 presents the net rate of return to private capital stock and the net rates of return to public infrastructure for the two extreme cases where $p_g = 0.89$ and $p_g = 1.34$, respectively. For the case of $p_g = 1.34$ [see the curve titled "social rate of return to public capital ($p_g = 1.34$)" in Figure 9], the conclusion regarding whether public infrastructure is over- or undersupplied is almost the same as that presented before. More specifically, over the subperiod from 1953 to 1973, the rate of return to public infrastructure is much higher than that to private capital, reflecting the shortage of public infrastructure in that period. Over the subperiod from 1974 to 2001, the rate of return to public infrastructure is roughly the same as that to private capital. As the value for p_g decreases to 0.89 [see the curve titled "social rate of return to public capital ($p_g = 0.89$)" in Figure 9], the time when the



FIGURE 9. Net social rate of return to public infrastructure and rate of return to private capital in total U.S. manufacturing industry.

rate of return to public infrastructure converges to that to private capital is delayed to the year of 1998. In particular, the rate of return to public infrastructure is higher than that to private capital over the period from 1953 to 1998, whereas the former is roughly the same as the latter afterward. Despite this difference, the temporal pattern exhibited by the net rate of return to public infrastructure in the case where $p_g = 0.89$ and that in the case where $p_g = 1.34$ are roughly the same; i.e., it first declines sharply and then slowly converges to that to private capital. Again, when the consumption sector and nonmanufacturing industries, together with the effects of the Tax Reform Act of 1986, are taken into account, public infrastructure is undersupplied in the post-1973 period in the case where $p_g = 1.34$ (or in the post-1998 period in the case where $p_g = 0.89$), for the same reason as discussed previously.

5.7. Robustness

As noted in Section 4.2, a possible problem with our estimate of the priceaugmenting AIM(2) cost function is endogeneity. More specifically, the input prices and public infrastructure on the right-hand side of (29)–(32) may not be exogenous. For the input prices, their possible endogeneity can be a result of the monopsonistic power exerted by some of the firms in the industry. As for public infrastructure, we conducted a Granger causality test to examine the direction of the relationship between private production (i.e., output) and public infrastructure for U.S. total manufacturing. Our test results show that the causation is unclear,



FIGURE 10. Cost elasticity with respect to public infrastructure in total U.S. manufacturing industry from the 2SLS model.

in that public infrastructure Granger-causing private production as well as private production Granger-causing public infrastructure. Thus, in this subsection we examine the sensitivity of our results (regarding cost elasticities, output elasticities, productivity decomposition, and rates of return) to the use of the two-stage least squares (2SLS) approach.

In using the 2SLS approach, our instrument set for input prices, p_i (i = 1, 2, 3), includes $p_i(-1)$ (where -1 indicates that p_i is lagged by one period), $p_i(-2)$, $p_y(-1)$ (where p_y is the output price), and $p_y(-2)$. The use of lagged input and output prices as instrumental variables is based on the argument that they are predetermined and thus are more likely to be exogenous. For the case of public infrastructure, our instrument set includes g(-1) (where g is public infrastructure as defined previously), g(-2), pop(-1) (where pop indicates population), and pop(-2). The lagged values of both public infrastructure and population are widely used in this literature to instrument public infrastructure. See, for example, Munnell (1992) and Easterly and Rebelo (1993).

The empirical results for total U.S. manufacturing, obtained using the 2SLS approach, are summarized in Figures 10–15. A comparison between Figures 1, 3, 4, 5, 6, and 7 and Figures 10–15 reveals that the major conclusions reached in the previous four subsections are still valid, although we notice some slight changes in the magnitude of the cost elasticities, output elasticities, productivity estimates, and rates of return to public infrastructure. First, as shown in Figure 10, the temporal pattern of the cost elasticity with respect to public infrastructure, estimated from the 2SLS model, is very similar to that from the Zellner method.



FIGURE 11. Input-specific elasticities with respect to public infrastructure in total U.S. manufacturing industry from the 2SLS model.



FIGURE 12. Contribution of input factors to total cost elasticity in total U.S. manufacturing industry from the 2SLS model.



FIGURE 13. Output elasticities in total U.S. manufacturing industry from the 2SLS model.



FIGURE 14. Productivity growth decomposition in total U.S. manufacturing industry from the 2SLS model.



FIGURE 15. Social rate of return to public infrastructure and rate of return to private capital in total U.S. manufacturing industry from the 2SLS model.

In particular, the cost elasticity with respect to public infrastructure, estimated from the 2SLS model, is also negative throughout the sample period with an average of -0.1601, suggesting that public infrastructure is productive in terms of cost reduction. In addition, as with the Zellner method, the cost elasticity with respect to public infrastructure, estimated from the 2SLS model, also shows a break in the early 1970s. More specifically, the estimates of the cost elasticity are less negative over time, decrease rapidly from 0.2350 in 1955 to 0.1426 in 1975 in absolute terms, and then stabilize at around 0.1400 in absolute terms in the second subperiod from 1976 to 2001. Moreover, public capital also mostly increases the efficiency of physical capital—see Figures 11 and $12.^2$

Second, as with the Zellner method, although the output elasticity with respect to public infrastructure and those with respect to private capital are comparable to each other in terms of magnitude, they exhibit very different temporal patterns. As shown in Figure 13, while the output elasticity of private capital is relatively stable at around 0.0879 over the sample period, that of public infrastructure declined steadily from 0.2350 to 0.1284 over the sample period, suggesting that public infrastructure has become less productive over time.

Third, like the results obtained from the Zellner method, public infrastructure is still a significant contributor to productivity growth in the first subperiod, but only a limited contributor in the second subperiod. More specifically, as shown in Figure 14, on average public infrastructure makes an impressive contribution to total productivity growth (50.0%) in the first subperiod (i.e., the pre-1976 period), but a much smaller contribution (19.2%) in the second subperiod (i.e., the post-1976 period).

Fourth, our conclusion that public capital is undersupplied still holds. More specifically, as shown in Figure 15, for the first subperiod the net rate of return to public infrastructure averages 27.03%, much higher than that to private capital (9.58%) over the same period. This implies that as with the Zellner method, public infrastructure is undersupplied in the first subperiod. For the second subperiod, the former (15.31%) is still higher than the latter (8.00%), implying that public infrastructure is undersupplied in the second subperiod too. This conclusion can be further strengthened if we take into account the effect of the Tax Reform Act of 1986 on p_g and the spillover effects of public infrastructure in the consumption sector and nonmanufacturing industries, as they will all cause the net rate of return to public infrastructure to be even higher than that shown in Figure 15.

5.8. Discussion

As noted from the results from both the Zellner and 2SLS approaches, a most salient one is that there is a "break" in the mid-1970s for all the estimates of the cost elasticities, output elasticities, productivity growth, and social rate of return to public infrastructure. For simplicity, in what follows we concentrate on the results from the Zellner method. More specifically, the estimates of the cost elasticity with respect to public infrastructure decrease rapidly in the first subperiod (from 1953) to 1973) in absolute terms, but then flatten out in the second subperiod (from 1974 to 2001). Similarly, the estimates of the output elasticity with respect to public infrastructure decrease rapidly in the first subperiod, but then become flat in the second subperiod. In terms of its contribution to total productivity growth, public infrastructure plays an important role in the first subperiod and only a minor one in the second subperiod. In addition, the rate of return to public infrastructure falls rapidly in the first subperiod, but then it flattens out in the second subperiod. Because this break is of particular importance in answering the aforementioned public policy questions, it is worth discussing in more detail the causes of the break.

From an empirical perspective, the break is actually caused by the negative sign of the estimated curvature parameters, λ_i (i = 1, 2, 3). As noted in Section 3.2, one of the powerful features of the Box–Cox function, which is used to model the effect of public infrastructure on the efficiency of private inputs in (16), is that it includes as special cases all the common "linear in the parameters" forms: linear, reciprocal (hyperbolic), log, polynomial, etc., depending on the estimated value of the curvature λ_i . The resulting functional form is an outcome of the estimation process; the researcher need not specify the exact functional representation but only a parametric family of functions, and, via the estimation process, the "best fitting" functional form is chosen. In other words, the Box–Cox function enables us to let the data determine the best functional form for the effect of public infrastructure on the efficiency of private inputs, which is a powerful attribute of our price-augmenting AIM cost function.

In our particular case, all three estimated curvature parameters for public infrastructure, λ_i (i = 1, 2, 3), are negative. More specifically, as shown in Table 2, $\lambda_1 = -0.4220$, $\lambda_2 = -2.9598$, and $\lambda_3 = -6.0000$. With γ_i (i = 1, 2, 3) being positive (see Table 2), and ln g and ln g₀ also being positive (because both g and g₀ are greater than one), the negative sign of λ_i (i = 1, 2, 3) implies that the ability of public infrastructure to increase the efficiency level of private inputs has declined over the sample period. More formally, taking the derivative of (18) with respect to ln g yields

$$\frac{\partial^2 \ln C}{\partial^2 \ln g} = -\sum_{i=1}^n s_i \left[\gamma_i \left(\lambda_i - 1 \right) \left(\frac{\ln g}{\ln g_0} \right)^{\lambda_i - 2} \right] > 0.$$
(36)

Because $\partial^2 C / \partial^2 g$ has the same sign as $\partial^2 \ln C / \partial^2 \ln g$ and $\partial \ln C / \partial \ln g < 0$ (see (18)), (36) implies that as public infrastructure increases over time, an additional one-unit increase in public infrastructure can still lead to a reduction in total cost, but at a lower rate. Similarly, we can show that

$$\frac{\partial^2 \ln y}{\partial^2 \ln g} = \sum_{i=1}^n s_i \left[\gamma_i \left(\lambda_i - 1 \right) \left(\frac{\ln g}{\ln g_0} \right)^{\lambda_i - 2} \right] < 0, \tag{37}$$

where y is output, implying that as public infrastructure increases over time, an additional one-unit increase in public infrastructure can still lead to an increase in output, but at a lower rate. In other words, both (36) and (37) imply that public infrastructure becomes less productive over time.

Moreover, when the ratio of $\ln g$ to $\ln g_0$ becomes high enough at a certain point in time, it is easy to see that the negative sign of λ_i (i = 1, 2, 3) implies that $\partial^2 \ln C/\partial^2 \ln g$ in (36) and $\partial^2 \ln y/\partial^2 \ln g$ in (37) tend to be zero, implying that an additional one-unit increase in public infrastructure leads to a rather constant increase in output or a rather constant decrease in cost. This means that cost elasticities, output elasticities, and the social rate of return to public infrastructure flatten out in the subsequent period. Thus the point of time at which an additional one-unit increase in public infrastructure starts resulting in a rather constant decrease in cost elasticities/output elasticities/the social rate of return to public infrastructure in absolute terms forms a break; i.e., prior to this point, cost elasticities (output elasticities, and the social rate of return) decrease in absolute terms, and after it, they flatten out. In our particular case, the break happens at around 1973–1974.

Having discussed the cause of the break from an empirical perspective, we now discuss the possible theoretical reason behind the break. Although our framework does not allow us to theoretically identify the causes behind the break, we believe that it is caused by the "network effect," discussed in Fernald (1999). This is because, as in Fernald (1999), the core public infrastructure used in this paper is composed mainly of highways (around 60%), which are well documented to

have a network effect. See, for example, Liebowitz and Margolis (1994), the U.S. Department of Transportation (1996), and Fernald (1999). More specifically, after applying a growth accounting method to 29 sectors in the U.S. economy, Fernald (1999, p. 621) concludes that

the massive road-building of the 1950's and 1960's offered a one-time boost to the level of productivity, rather than a path to continuing rapid growth in productivity. This conclusion—that roads were exceptionally productive before 1973 but not exceptionally productive at the margin—is consistent with simple network argument. In particular, building an interstate network might be very productive; building a second network may not.

Clearly, our results from applying the price-augmenting AIM cost function, proposed in this paper, to the data on the U.S. manufacturing industry are perfectly consistent with this network argument.

6. CONCLUSION

In this paper, for the first time in the literature, we develop a price-augmenting AIM to investigate the effects of public infrastructure on the productivity of U.S. manufacturing industry. In doing so, we make a distinction between the productivity effect and the production factor effect of public infrastructure. This distinction allows us to investigate the more interesting productivity effect by incorporating public infrastructure into the AIM cost function through the efficiency index. This is in contrast to previous studies where public infrastructure is incorporated into cost functions as a fixed input, so that they are incapable of disentangling the productivity effect from the production factor effect of public infrastructure. The globally flexible AIM cost function is chosen because it is capable of approximating the underlying cost function at every point in the function's domain by increasing the order of the expansion, whereas most of the locally flexible functional forms (i.e., generalized Leontief, translog, and normalized quadratic) theoretically can attain flexibility only at a single point or in an infinitesimally small region. Further, the specification of the efficiency index as an exponential Box-Cox function enables us to investigate the temporal pattern of the effects of core public infrastructure with better insights.

We then analyze the effects of public infrastructure using four measures: the cost elasticity, the output elasticity, the contribution of public infrastructure to total factor productivity, and the social rate of return. All our results point to the decline in the spillover effects of core public infrastructure on the performance of manufacturing industry in the United States. In particular, our estimates of cost elasticities suggest that although public infrastructure offers significant benefits in terms of cost reduction in the pre-1973 period, it cannot offer the same benefits at the margin. Although our estimates of output elasticities suggest that the average output elasticity of core public infrastructure (0.0993) is by no means small, compared to that of capital, labor, and materials (0.1750, 0.4290, and 0.3962, respectively), they also indicate that core public infrastructure has become less

productive relative to private inputs over the sample period. More importantly, our analysis of the contribution of core public infrastructure to total factor productivity further confirms the finding in Fernald (1999) that the massive core infrastructure building of the 1950s, 1960s, and early 1970s offered a one-time increase in the level of productivity, rather than a continuing path to productivity growth. Finally, our results also indicate that the social rate of return to core infrastructure was high during the 1950s and the 1960s, but declined considerably over the sample period. After the 1980s, the rates of return to core public infrastructure and private sector capital seem to have converged. The results from our 2SLS approach are roughly the same, despite some small changes in magnitude.

We note that these results are obtained by using a two-digit classification of U.S. manufacturing industries. In our future research, we will use a four-digit or a six-digit classification of U.S. manufacturing industries, which we believe is more informative. This is because a lot of useful information contained in more disaggregated data may be lost when the data are aggregated. Thus we expect that the results obtained by using a four-digit or a six-digit classification will show more variation in terms of the effects of public infrastructure across industries.

NOTES

1. For the fact that constant returns to scale has been built in the KLEMS data, see the Bureau of Labor Statistics (BLS) Web site: http://www.bls.gov/mfp/mprover.htm#On_the_Internet. It was also confirmed by our personal communication with Randy Kinoshita, the economist in charge of the KLEMS data at the BLS.

2. Notice that because $p_i(-2)$ and g(-2) are used in the 2SLS model, the first two observations are lost. Also, because of the use of the lagged values of input prices and public infrastructure as instruments, the break is delayed to 1976.

REFERENCES

Aschauer, D. (1989) Is public expenditure productive? *Journal of Monetary Economics* 23, 177–200. Ballard, C.L. (1990) Marginal efficiency cost calculations: Differential analysis vs. balanced budget

analysis. Journal of Public Economics 41, 263-276.

- Barnett, W.A., J. Geweke, and M. Wolfe (1991) Semi-nonparametric Bayesian estimation of the asymptotically ideal production model. *Journal of Econometrics* 49, 5–50.
- Barnett, W.A. and A. Serletis (2008) Consumer preferences and demand systems. Journal of Econometrics 147, 210–224.

Berndt, E. and B. Hansson (1992) Measuring the contribution of public infrastructure in Sweden. *Scandinavian Journal of Economics* 94, 151–168.

Browning, E.K. (1987) On the marginal welfare cost of taxation. *American Economic Review* 77, 11–23.

Bureau of Economic Analysis (June 1999) Note on Rates of Return for Domestic Nonfinancial Corporations, 1960–98. Survey of Current Business (http://www.bea.gov/scb/account_articles/ national/0699cba/maintext.htm).

Burguete, J.F., A.R. Gallant, and G. Souza (1982) On unification of the asymptotic theory of nonlinear econometric models. *Econometric Reviews* 1, 151–190.

Caves, D. and L.R. Christensen (1980) Global properties of flexible functional forms. *American Economic Review* 70, 422–432.

- Caves, D.W., L.R. Christensen, and W.E. Diewert (1982) The economic theory of index numbers and the measurement of input, output, and productivity. *Econometrica* 50, 1393–1414.
- Christensen, L., D.W. Jorgenson, and L.J. Lau (1975) Transcendental logarithmic utility functions. *American Economic Review* 65, 367–364.
- Diewert, W.E. (1971) An application of the Shephard duality theorem: A generalized Leontief production function. *Journal of Political Economy* 79, 481–507.
- Diewert, W.E. and K.J. Fox (2008) On the estimation of returns to scale, technical progress and monopolistic markups. *Journal of Econometrics* 145, 174–193.
- Diewert, W.E. and T.J. Wales (1987) Flexible functional forms and global curvature conditions. *Econometrica* 55, 43–68.
- Diewert, W.E. and T.J. Wales (1992) Quadratic spline models for producer's supply and demand functions. *International Economic Review* 33, 705–722.
- Duggal, V.G., C. Saltzman, and L.R. Klein (1999) Infrastructure and productivity: A nonlinear approach. *Journal of Econometrics* 92, 47–74.
- Duggal, V.G., C. Saltzman, and L.R. Klein (2007) Infrastructure and productivity: An extension to private infrastructure and its productivity. *Journal of Econometrics* 140, 485–502.
- Easterly, W. and S. Rebelo (1993) Fiscal policy and economic growth: An empirical investigation. *Journal of Monetary Economics* 32, 417–458.
- Evans, P. and G. Karras (1994) Are government activities productive? Evidence from a panel of U.S. states. *Review of Economics and Statistics* 76, 1–11.
- Feng, G. and A. Serletis (2008) Productivity trends in U.S. manufacturing: Evidence from the NQ and AIM cost functions. *Journal of Econometrics* 142, 281–311.
- Feng, G. and A. Serletis (2009) Efficiency and productivity of the US banking industry, 1998–2005: Evidence from the Fourier frontier function satisfying global regularity condition. *Journal of Applied Econometrics* 24, 105–138.
- Fernald, J.G. (1999) Roads to prosperity? Assessing the link between public capital and productivity. *American Economic Review* 89, 619–638.
- Gallant, A.R. (1975) Seemingly unrelated nonlinear regressions. Journal of Econometrics 3, 35–50.
- Gallant, A.R. (1982) Unbiased determination of production technology. *Journal of Econometrics* 20, 285–323.
- Gallant, A.R. and G. Golub (1984) Imposing curvature restrictions on flexible functional forms. *Journal of Econometrics* 26, 295–321.
- Gramlich, E.M. (1994) Infrastructure investment: A review essay. *Journal of Economic Literature* 32, 1176–1196.
- Guilkey, D., C. Lovell, and R. Sickles (1983) A comparison of the performance of three flexible functional forms. *International Economic Review* 24, 591–616.
- Hashimzade, N. and G. Myles (2010) Growth and public infrastructure. *Macroeconomic Dynamics* 14, 258–274.
- Holtz-Eakin, D. (1994) Public-sector capital and the productivity puzzle. *Review of Economics and Statistics* 76, 12–21.
- Hulten, C.R. and R.M. Schwab (1991) Public capital formation and the growth of regional manufacturing industries. *National Tax Journal* 44, 121–134.
- Jacobs, B. (2009) The Marginal Cost of Public Funds and Optimal Second-Best Policy Rules. Mimeo, Erasmus University, Rotterdam.

Jorgenson, D.W. (1996) Investment: Tax Policy and the Cost of Capital. Cambridge, MA: MIT Press.

- Jorgenson, D.W. and K. Yun (1990) Tax reform and U.S. economic growth. *Journal of Political Economy* 88, 151–93.
- Kaizuka, K. (1965) Public goods and decentralization of production. *Review of Economics and Statistics* 47, 118–120.
- Kemmerling, A. and A. Stephan (2002) The contribution of local public infrastructure to private productivity and its political economy: Evidence from a panel of large German cities. *Public Choice* 113, 403–424.

- Kleven, H. and C. Kreiner (2003) The Marginal Cost of Public Funds in OECD Countries. Hours of Work Versus Labor Force Participation. CESifo Working Paper 935.
- Kohli, U. (1990) Growth accounting in the open economy: Parametric and nonparametric estimates. Journal of Economic and Social Measurement 16, 125–136.
- Lau, L.J. (1978) Testing and imposing monotonicity, convexity, and quasi-convexity constraints. In M. Fuss and D. McFadden (eds.), *Production Economics: A Dual Approach to Theory and Applications*, vol. 1, pp. 409–453. Amsterdam: North Holland.
- Liebowitz, S.J. and S.E. Margolis (1994) Network externality: An uncommon tragedy. Journal of Economic Perspective 8, 133–150.
- Lucas, R.E. Jr. (1988) On the mechanics of economic development. *Journal of Monetary Economics* 22, 3–42.
- Meade, J.E. (1952) External economies and diseconomies in a competitive situation. *Economic Journal* 62, 54–67.
- Morey, E.R. (1986) An introduction to checking, testing, and imposing curvature properties: The true function and the estimated function. *Canadian Journal of Economics* 19, 207–235.
- Morrison, C.J. and A.E. Schwartz (1996) State infrastructure and productive performance. American Economic Review 86, 1095–1111.
- Morrison, C.J. and D.S. Siegel (1999) Scale economies and industry agglomeration externalities: A dynamic cost function approach. *American Economic Review* 89, 272–290.
- Munnell, A.H. (1990) How does public infrastructure affect regional economic performance? New England Economic Review (September/October), 11–32.
- Munnell, A.H. (1992) Policy watch: Infrastructure investment and economic growth. Journal of Economic Perspectives 6, 189–198.
- Nadiri, M.I. and T.P. Mamuneas (1994) The effects of public infrastructure and R&D capital on the cost structure and performance of U.S. manufacturing industries. *Review of Economics and Statistics* 76, 22–37.
- Nadiri, M.I. and T.P. Mamuneas (1998) Contribution of Highway Capital to Output and Productivity Growth in the U.S. Economy and Industries. Working Paper, U.S. Department of Transportation.
- Romer, P. (1986) Increasing returns and long-run growth. *Journal of Political Economy* 94, 1002– 1037.
- Romp, W. and J. de Haan (2007) Public capital and economic growth: A critical survey. *Perspektiven der Wirtschaftspolitik* 8, 6–52.
- Samuelson, P.A. (1954) The theory of public expenditure. *Review of Economics and Statistics* 36, 386–389.
- Serletis, A. and G. Feng (2010) Semi-nonparametric estimates of currency substitution between the Canadian dollar and the U.S. dollar. *Macroeconomic Dynamics* 14, 29–55.
- Serletis, A. and A. Shahmoradi (2005) Semi-nonparametric estimates of the demand for money in the United States. *Macroeconomic Dynamics* 9, 542–559.
- Serletis, A. and A. Shahmoradi (2008) Note on finite approximations of the asymptotically Ideal model. *Macroeconomic Dynamics* 12, 579–590.
- Sturm, J.E. and J. de Haan (1995) Is public expenditure really productive? *Economic Modelling* 12, 60–72.
- Sturm, J.E., G.H. Kuper, and J. de Haan (1998) Modelling government investment and economic growth on a macro level: A review. In S. Brakman, H. Vanes, and S.K. Kuipers (eds.), *Market Behaviour and Macroeconomic Modelling*, pp. 359–406. London: MacMillan.
- Tatom, J.A. (1991) Public capital and private sector performance. *Federal Reserve Bank of St. Louis Review* 73, 3–15.
- Thomsen, T. (2000) Short cuts to dynamic factor demand modelling. *Journal of Econometrics* 97, 1–23.
- U.S. Department of Transportation (1996) Productivity and Highway Network (www.fhwa.dot.gov/ policy/otps/060320b/index.htm).

APPENDIX A

From equation (8) we get

$$\epsilon_{ct} = \frac{\partial \ln C}{\partial t} = \frac{1}{C} \frac{\partial C}{\partial t} = \frac{1}{C} \left(\sum_{i=1}^{n} \frac{\partial C}{\partial P_n} \frac{\partial P_n}{\partial t} \right).$$

Using Shephard's lemma, equation (10), this yields

$$\epsilon_{ct} = \frac{\partial \ln C}{\partial t} = \frac{1}{C} \left(\sum_{i=1}^{n} X_n \frac{\partial P_n}{\partial t} \right) = \frac{1}{C} \left(\sum_{i=1}^{n} P_n X_n \frac{1}{P_n} \frac{\partial P_n}{\partial t} \right)$$
$$= \sum_{i=1}^{n} \frac{P_n X_n}{C} \frac{\partial \ln P_n}{\partial t} = \sum_{i=1}^{n} s_i \frac{\partial \ln P_n}{\partial t}$$
$$= \sum_{i=1}^{n} s_i \frac{\partial (\ln P_n - \ln A_n)}{\partial t},$$

and because p_n is exogenous and not a function of t, this reduces to equation (11).

APPENDIX B

From equation (8) or (12), we get

$$\frac{\partial \ln C}{\partial \ln g} = -\sum_{i=1}^n s_i \frac{\partial (\ln A_n)}{\partial \ln g}.$$

Moreover, equation (16) can be written as

$$\ln A_i = \frac{\vartheta_i t_0}{\delta_i} \left[\left(\frac{t}{t_0} \right)^{\delta_i} - 1 \right] + \frac{\gamma_i \ln g_0}{\lambda_i} \left[\left(\frac{\ln g}{\ln g_0} \right)^{\lambda_i} - 1 \right],$$

so that

$$\frac{\partial (\ln A_i)}{\partial \ln g} = \frac{\gamma_i \ln g_0}{\lambda_i} \lambda_i \left(\frac{\ln g}{\ln g_0}\right)^{\lambda_i - 1} \frac{1}{\ln g_0} = \gamma_i \left(\frac{\ln g}{\ln g_0}\right)^{\lambda_i - 1}.$$

Thus,

$$\frac{\partial \ln C}{\partial \ln g} = -\sum_{i=1}^n s_i \frac{\partial (\ln A_n)}{\partial \ln g} = -\sum_{i=1}^n s_i \gamma_i \left(\frac{\ln g}{\ln g_0}\right)^{\lambda_i - 1}$$