Undergraduate Economic Review

Volume 12 | Issue 1 Article 4

2015

Cost - Benefit Analysis of the German High Speed Rail Network

Martin Dorciak
Lewis & Clark College, mdorciak@lclark.edu

Cost - Benefit Analysis of the German High Speed Rail Network

Abstract

This study undertakes a cost-benefit analysis of the German railway market looking specifically at the effects of high-speed rail development on railway passenger subsidies. Using OLS regression analysis, I estimate a demand curve for the German railway network at the route level; this is combined with cost curve estimates to yield a required subsidy for rail development assuming a natural monopoly market structure. I find that an increase in demand as a result of the introduction of high-speed rail technology causes a 23.9% decrease in required rail subsidies.

Keywords

Consumer Surplus, Cost Benefit, Cost Effective, Elasticity of Demand, Railways, Regional Transportation, Transportation

1 Introduction

Medium distance travel offers the widest array of choices for travellers in terms of modes of transport. Within the past 30 years, the advent of high-speed rail technology has once again enabled rail to be a player in medium distance travel of 200 to 500 miles, enabling railways to compete with the currently more appealing modes of cars and planes by reducing the costs and travel time associated with rail travel. The current developments in railway transport have allowed rail to increase its modal share relative to the other two modes, where modal share is the percentage of passengers that utilize one mode relative to all available options. The main motivation of this study is to estimate the impact of high-speed rail (henceforth referred to as HSR) technology on the demand for rail transport and compare this with the costs of implementing HSR. Then I comment on the viability of HSR as an alternative option to air and vehicular travel. I will look specifically at the case of Germany and comment on the effectiveness of the German HSR experience in terms of costs and benefits.

The roots of HSR analysis come from transport economics and valuations of travel time and fare prices and how they affect consumer demand. Given a choice of modes on a given route, consumers choose the mode that minimizes the total trip cost, which equals the sum of the fare, the cost of preliminary transit (transit to a station/airport), wait times, and the cost of the actual travel time itself. In the specific valuation of transportation demand, the primary method is that of an aggregate transportation model, where "the demand for a given [mode in] the travel market is explained as a function of variables that describe the product or its consumers" (Small, 1992: 6). This derivation of transport demand can be focused even further as in Domencich and Krafts (1970) approach of direct demand modeling, where the dataset to be analyzed is comprised of pairs of locations forming various routes in a city or region from which travel statistics such as ridership figures, fare prices, and regional statistics are regressed to determine an explicit demand curve for the specified region. I will employ this particular method to estimate demand, focusing on a route level analysis of the German railway network while also differentiating between routes that have HSR service and routes that do not.

On the supply side, looking at costs and market structures further reveal key motivations for this study. The predominant market structure for railroad systems is that of a natural monopoly, where large economies of scale and prohibitively high startup costs limit the number of firms to one. Even then however, the suppliers of rail service typically require subsidization to prevent travel costs from being prohibitively high for the consumer. Mohring (1972) provides a theoretical framework for optimal subsidization levels in public transport and posits that if the average riders opportunity cost of travel time is below the marginal cost of operating a transport system, then to make marginal cost pricing viable transport providers would have to subsidize fares to a level that equals the opportunity cost of the passengers time. This can also be intuited in the case of the monopolistic railway networks. But rather than subsidizing a consumers time, production costs must be subsidized to make monopoly pricing viable.

This issue is certainly relevant to current political debates, with developments proposed in California and Texas funded by an 8 billion dollar allotment from the 2009 American Recovery and Reinvestment Act (Peterman, 2009). Currently, several Central and Western European countries have developed HSR networks with varying degrees of coverage that offer travellers overall travel times which are shorter than flying or driving at comparable or lower prices for medium distances of 100 to 500 miles with almost exclusive provision of services through state owned monopolies who in turn subsidize construction costs and fares with government money. This is indicative of the motivations behind my analysis, as it is in the interest of national governments to pursue HSR development policies; I aim to comment on the viability of these developments.

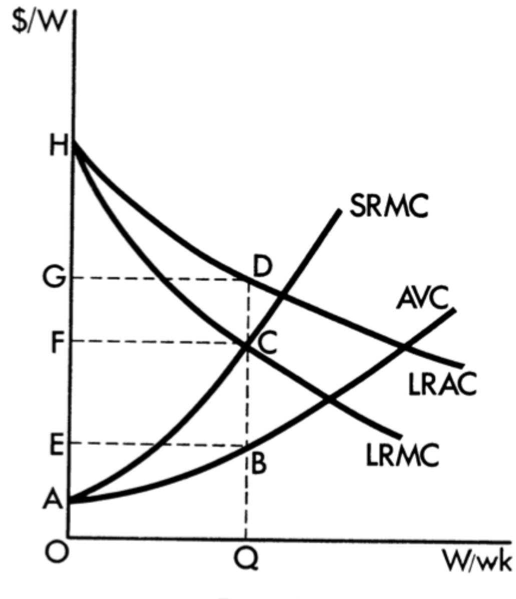
The core of my analysis will focus on the derivation of the demand and cost curves through a regression analysis of German rail statistics, specifically focusing on observations of city pairs where the quantity of rail transport demanded is represented by the route ridership and price is represented by the total cost to passengers on the given route. Additional variables include a dummy variable indicating the presence of HSR on a given route to differentiate between base level rail demand and high speed rail demand, average population density of the departure and arrival destinations, average GDP per capita of departure and arrival destinations, and travel times and fare prices of the primary competitor in that of air travel. The analysis combines this market demand curve (equal to marginal benefit) with costs incurred by the monopolist firm to determine the socially optimal level of HSR utilization for the German market and comment on discrepancies between optimal and actual levels of ridership and commensurate subsidy levels.

The paper will proceed as follows. Section 2 reviews the literature in the field looking at general theories of transportation economics as they pertain to institutionally provided transport and subsidization of said transport, and specific applications of HSR analyses estimating costs, demand, and societal benefits. Section 3 presents the economic framework typical to the state monopolist market present in Germany and equates curves in the ideal model with their function in my analysis. Section 4 presents the data and regression analysis to derive the demand curve and the specification of cost curves and section 5 represents these graphically and determines the optimal level of utilization and compares this with observed levels.

2 Literature Review

Mohring (1972) develops a model that illustrates the role of subsidization in public transport systems and then applies his model to travel statistics for city busses in the Minneapolis St. Paul area. Mohring provides a general base for analysis of transport subsidization, and additionally provides specification of several key cost variables. First, Mohring notes how transport pricing deviates from traditional price theory models, where travellers also provide an input into the final cost calculations in that of their travel time. This makes the total cost to consumers of utilizing transport the fare that is charged in addition to the opportunity cost of the journey for the consumer. The model presented by Mohring is presented below on the folowing page in Figure 1.

Figure 1: Bus Transport Subsidies



The commodity detailed in the graph is bus rides; as such, cost is represented as dollars per bus ride and quantity as bus rides per week, with curves representing short run and long run marginal costs, long run average costs, and average variable costs. Mohring assumes this form without alteration [describes] bus operations that are subject to increasing returns and can be applied practically to metropolitan transit systems (Mohring, 592, 1972). Additionally, a demand curve is implied as intersecting marginal costs at point C. This model reconciles both the inputs of the consumer in that of travel time and the producer in that of the bus system and widgets are transformed into journeys. Mohring assumes a consumer provides time inputs valued at E for point B on the average variable cost curve, noting the curves representation of consumer valuation of time. Then, Mohring notes fare price at the level F, where demand intersects marginal costs. This would generate what Mohring refers to as quasi-rent for bus services of EBCF, but ultimately this falls short of the costs of fixed inputs, and requires a subsidy of FCDG to meet those input prices. I will use a variant of this model, representing a natural monopoly structure to estimate required subsidies to supplement fare revenues, something I will elaborate on in section 3.

Empirical studies of HSR focus predominantly on route level and network (national) level analyses. Couto and Graham (2007) conduct one such analysis, specifically looking at the impact of HSR service on quality and speed of travel and the commensurate affect on their derivation of a rail demand function at national levels. The study first posits a basic demand relationship where a demand for railroad services is expressed as a function of fare price, level of income, and other factors such as the presence of HSR (represented as a dummy), geographical and economic conditions such as regional incomes and city size of destinations, and prices of alternate modes of travel (Cuoto and Graham, 114, 2007). To estimate their model, Cuoto and Graham use a log-linear form 2SLS model to estimate coefficients that are themselves the respective [demand function] elasticities (Cuoto and Graham, 120, 2007). They regress y passenger-kilometers per kilometer of network length, (roughly an indicator of network utilization relative to the size of the network) on four dummy variables representing the introduction and use of two different HSR technologies (conventional HSR, and tilting HSR which runs on existing lines).

Additional variables include fare prices, alternate transport prices, and national geographical and economic indicators. Cuoto and Graham chose the 2SLS model due to the structure of their variable representing railway demand. The demand variable took the form of passenger kilometers per kilometer of network length, and the 2SLS form was necessary due to the endogeneity of passenger kilometers (a measure of total kilometers travelled by all passengers) in their regression. Their regression results indicate a price elasticity of demand of -0.22, which is in line with estimates of price elasticity from similar studies by Fitzroy and Smith (1995,1998) and McGeehan (1984). They further find that conventional high-speed technology increases railway demand overall at a national level, which coincides with Fowkes and Nash (1991) who concluded that regardless of speed increases, demand for rail travel will raise with the presence of a national HSR network. In turn this suggests a 9% increase in passenger demand for railway travel given the existence of a HSR network. The variable for tilting high-speed technology was not found to be statistically significant.

Ultimately Cuoto and Graham conclude that if railway development is to be supported, then HSR technology is a viable method to increase rail demand. The study concludes with a brief statement as to the benefits of HSR investment in that both tilting and conventional HSR can result in increased demand with tilting technology requiring a much lower initial investment (due to utilization of exiting networks) but only increasing demand due to increased frequency, and conventional technology requiring a much higher investment but resulting in greater overall demand. This is key to my study and is a central assumption that I will test for the German market. Rather, that HSR service inherently increases demand on a given route, as opposed to a network. The magnitude increase estimated by Cuoto and

Graham will also be useful to compare my conclusions to and to note the difference in magnitudes between route level and network level analyses.

Behrens and Pels (2010) look specifically at intermodal competition on the London Paris corridor where they work to determine passenger choices between HSR and air travel from passenger preferences as observed surveys and explicit ridership data. Behrens and Pels utilize nested logit and mixed logit models to explain passenger choice on the London Paris route. Their study draws from past work utilizing logit regressions to determine passenger behavior, and draw their key variables from these studies in that of fares, accessibility (to transport), frequency (of offered journeys on a route), and reason for travel (leisure or business). Specifically, Behrens and Pels use similar characteristics in that of travel costs and times for rail and substitutes, frequency, delay times, and personal statistics; they also differentiate regressions between leisure travellers and business travellers under the assumption of different values placed on particular characteristics.

The model specified takes two forms, first a nested form where choice of mode is nested first by type (air or rail) and then by specific route (arrival and departure airport pairs and different HSR services). The second form merely groups all the possible routes into one group and assumes consumers choose mode and route in one step. Their results concluded that HSR is a viable competitor for air travel on the specified route. Business travelers in particular were found to choose HSR due to their frequency of service over air travel. Conversely, it was found that leisure passengers tended to substitute into HSR in the face of rising airfare costs, and the relative stability of HSR fares. The authors additionally note that several of the flights (by various airlines) between the cities that they used in the model have since been discontinued, showing the practical effects of passengers choosing HSR over air travel on routes where airlines encounter strong competition from rail. Both this study and Cuoto and Graham provide insight into my own study regarding how to approach the functional form of my demand regression as well as detailing key variables of interest.

On the supply side, Levinson, Mathieu, Gillen, and Kanafani (1997) conduct a theoretical study to estimate the full costs of developing a high speed rail corridor in the state of California based on past observations and economic theory. Their primary motivation lies in determining the private and social costs of providing an intercity HSR network and to determine what these costs imply when looking at HSR investment as opposed to air travel and highway infrastructure investment. One of the key issues with such an analysis as noted by the authors is the differentiation between costs paid by the users of the transport mode and costs paid by others. Their taxonomy for full cost to society (FC) includes numerous variables but chiefly among them are infrastructure costs for construction (IC) and maintenance (MC) of the rail network, carrier capital and operating costs incurred by operators of rail service for the purchase of vehicles and other items (CC), user money costs (fares and fees, FF), user travel time costs (opportunity cost of passengers, represented by TC), user delay costs (opportunity cost of delays, DC), and social costs incurred by people exogenous to the system in that of emissions and noise pollution (SC). The full cost of high-speed rail development and operation on a given route is thus calculated as a sum of the aforementioned values and takes the following form.

$$FC = IC + MC + CC + FF + TC + DC + SC \tag{1}$$

The authors follow this specification with the caveat that "each cost is a function of various parameters and depends on the level of [ridership]" and devote the remainder of the paper to estimate these costs as functions of ridership (Levinson et al, 192, 1997). The infrastructure costs are stated as the costs of building the rail network and all that entails in terms of structures, rail, power, and earthworks; this is simply divided by the number of passengers utilizing the system to determine costs per passenger and leads to a negatively sloped average infrastructure cost as ridership increases. Carrier costs follow and are divided into two parts, namely carrier operating and capital costs. Both are determined by multiplying the cost per unit by the number of units in operation (where unit refers to a train) for operating expenses such as electricity and worker salaries, and capital expenses in that of the cost of the train. These costs will only vary with the number of units in service and assuming the rolling stock of rail companies remains constant in the short run, the marginal costs of accommodating additional travellers on a given train is essentially fixed. This study is unique among others in that the authors account fully for all costs incurred whereas other studies focus only on some sources of costs and quantify them differently (such as cost per passenger-kilometer, or passenger journey), Levinson et. al. instead focus on the full costs incurred by society. The remaining costs are noted as incurred by the consumer and the public in general. The authors posit theories for the remaining factors where noise and pollution can be interpreted by variations in housing price around HSR lines, but this section is generally speculative.

Levinson et. al. use these costs to determine the total cost per user of the system and compare it to the costs of air and vehicular travel. Additionally the authors continue to tie their estimates back to the proposed Californian HSR network. Their analysis and cost calculations are based on data from the European Union and Japan, however the authors acknowledge the differences and difficulty in comparisons of developments in other countries and the United States. The authors conclude by noting that it would be difficult to implement HSR in California without massive subsidization and even then so with higher subsidization than other countries. This is due to constraints on federal spending, the deregulated nature of air transport providing cheaper airfares (which was largely state directed in Europe and Japan during the advent of HSR), and the sheer geographical distances between Californian cities. However, the cost curve specifications in this model can be incorporated into the theory of Mohrings model and applied to the German case.

3 Economic Theory

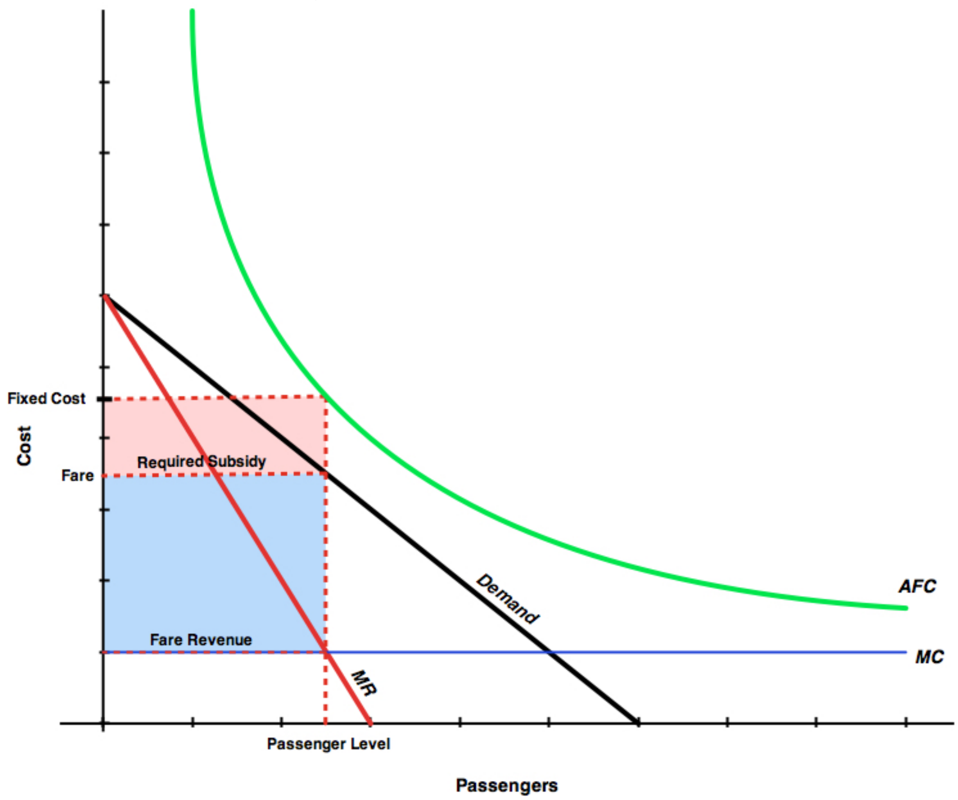
The essential theory of my analysis requires several preliminary model specifications and assumptions. First, it is important to note the prevailing market structure in HSR before pursuing further specificity. As is the case of most public transport goods, HSR networks are integrated into state owned railway companies and take the form of a natural monopoly (Mohring, 1972). Specifically, the railway industry has very high costs of production with the good in question being transport, but small and constant marginal costs; also featured are large economies of scale and a large infrastructure that is consolidated under the monopolist firm.

3.1 German Railway Market Structure

In the case of Germany, national transport and freight railroad is consolidated under Deutsche Bahn AG (henceforth referred to as DB), a private joint-stock company with 100% ownership by the German government (Deutsche Bahn). Dunn and Perl (1994) provide further information on the formation and structure of this network; DB stretches back to the Cold War and the reunification and integration of the two Germanies and how this effected the formation of a national railway system. This national disunity could explain the slow pace of the introduction of Germanys high-speed Intercity Express (ICE) train network. Additionally, the paper provides other details that are not mentioned in economic analyses that shed light on some added costs and rationale as to why the German network took the form that it did. The primary difference between German and other HSR according to the authors is the German utilization of standard previously existing track for both passenger and freight transport upgraded to handle HSR trains, a choice made to avoid prohibitively expensive costs of tunneling new HSR lines through mountains. This has several economic implications, namely the cost savings as opposed to constructing a new network, and the slightly slower travel times on German HSR relative to France and other nations.

All of this can be combined into a general model for a given railway route in Germany that I will use in my study, presented below in Figure 2. This model takes an adaptation of the general form presented by Mohring

Figure 2: Railway Route Model



(1972), where D = MC is the fair market fare price including time valuation, and MR = MC is the fare price charged. MC in this case is fixed due to the fixed nature of accommodating an additional passenger on rolling stock that has already been purchased, this is to say, assuming a fixed number of trains running a fixed number of routes, the cost of accommodating an additional passenger on a train is constant. Marginal revenue is derived from the demand curve by transforming demand into total revenue and then taking the derivative, from which I can determine the fare price charged on a given route by looking at the intersection of marginal revenue and marginal costs. As in a traditional monopoly, this determines the price and from that I can determine realized producer surplus as indicated above in Figure 2. This however falls short of the AFC curve, which here represents the infrastructure costs per rider incurred by DB in the construction of rail lines and modification of existing lines to accommodate high-speed trains. It is here that Mohrings theory comes in to play with DB being unable to cover total costs at any price or ridership level thus requiring subsidization to cover all costs. As follows, the distance between D and AC at the monopolist price level (the distance between Fare and Total Cost in Figure 2) is equal to the subsidy needed to cover all costs that are not incurred by riders at the monopolist price level. Additionally important to mention, due to the nature of my data this model compares costs and demand at the route level, where a route is defined as a path of travel between a city pair, rather than looking at the railway network in Germany as a whole.

This ultimately leads to the key point of interest in my study. A central assumption to my model is borrowed from Mohring in that of an average cost curve that is higher than demand for all quantities of passengers, and thus unlike a standard monopolist firm, railroad providers cannot charge a fare that will recoup all infrastructure and rolling stock costs and be low enough to maintain any passenger levels. In regards to HSR, Cuoto and Graham (2007) concluded that the presence of HSR led to an across the board increase in demand for rail service, in line with other studies. In the model this translates to a shift outward in the demand curve for a network with HSR service and as such leads to a higher ridership level and a lower required subsidy. This change in required subsidy level will be the key focus of my analysis, looking at if the introduction of HSR provides any meaningful magnitude change in the level. The next two subsections will focus on the specification of the functional form of the demand and cost curves used in my model.

3.2 Railway Demand

The central component of this study focuses on estimating a demand curve at the route level of the German railway network. I will estimate a linear demand specification for a given German route that takes the form of a traditional linear demand curve. The components that determine the curve are total cost (TC) of a rail journey to a consumer and quantity of rail travel demanded (D), which is represented by the number of rail passengers on a given route as used by De Rus and Inglada (1997) to estimate railway demand in Spain. As such, the relationship of interest is how the number of rail passengers responds to changes in the total cost of a given journey. The specific functional form of the regression utilized is as follows.

$$D = \beta_0 + \beta_1 TC + \beta_2 X_2 + ... + \beta_n X_n + \epsilon \tag{2}$$

This again represents the form assumed above, where railway demand is dependent on total cost of travel, with other explanatory variables also included. This regression yields an inverse function explaining the change in quantity demanded as a function of price. The demand function can be transformed,

$$TC = \frac{\beta_0}{\beta_1} + \frac{1}{\beta_1} D \qquad (3)$$

. This transformation allows my demand function to be represented graphically alongside the remaining curves in my model by standardizing axis variables and by retaining the relationships of interest yielded in the regression results.

3.3 Railway Costs

Costs in this model are relatively straightforward, and again are represented as the costs to consumers at the route level. The unique aspect of Germany and its usefulness in regards to this study is that of the nature of its railway system. The benefit lies in the fact that Germany uses the same track network for both normal and HSR applications, with both freight and passenger traffic also travelling on the same track. This is also a key determinant of which railway lines in Germany received HSR upgrades; while major population centers will be linked with high-speed service, the unification of freight and passenger traffic gives incentive to bring higher speed service to industrial centers that may not have much consumer demand for the service (Albalate and Germa, 2012). This would seem to further support the choice of Germany as an ideal location for this study, where passenger travel would seem to take a secondary role when it came to developing HSR lines to that of freight transport, thus leading to the conclusion that there may be some randomness in the distribution of the HSR treatment in my model (notwithstanding highly likely endogeneity issues between freight and passenger transport between cities).

Ultimately, this suggests however that a route level analysis and comparison of costs is viable as much of the German railroad network thus has a uniform average cost per kilometer of track. First, it is necessary to specify the average costs per passenger incurred by DB on infrastructure costs on a given route. I am assuming variable costs to be negligible in my model, as the two major cost groups (infrastructure and rolling stock) are both fixed for a given route, thus average costs would simply be fixed costs of infrastructure per route divided by passenger level. This simplification allows me to assemble a variable for route cost (RC) as the product of average cost per kilometer and route length. This cost per route must then be divided by the amortization period (AP) of the route development costs. This is then divided by passenger level (n) to determine the average fixed cost per passenger at a given passenger level. This is represented as follows.

$$AC = \frac{\frac{RC}{AP}}{n} \tag{4}$$

Additionally, since passenger level (n) is variable, this equation will take a downward convex form. This is intuitively sound and can even be equated to a long run average total cost exhibiting increasing returns to scale, given my assumption of negligible short run variable costs.

Marginal costs in my model are held fixed. This is rationalized due to the fixed capacity of the German rail system in that once trains are acquired, the cost of accommodating an additional passenger on a train is fixed up until the capacity of the system is reached and more trains must be purchased. Currently, the German rail system is comprised of refurbished trains from the 90s, newer low cost regional trains, and 67 high speed trains built by Siemens AG. Even in this regard, Germany is ideal for this study as rather than purchasing trains on an as-needed basis, DB purchased the trains in two major transactions, one in 2000 acquiring 50 Siemens Velaro ICE 3 Class trains with 30 years of maintenance and a second in 2008 with 17 Siemens Velaro D Class trains for international use based in Germany (Siemens, 2008). Thus to determine the marginal cost per passenger I take the total cost (TC) of all rolling stock acquired by Deutsche Bahn and divide it by the number of years in the period of amortization (AP) of the funds used to purchase the trains. This is then be divided by the annual maximum theoretical ridership (total train capacity multiplied by number of trains multiplied by journeys per day) (TR).

$$MC = \frac{\frac{TC}{AP}}{TR} \tag{5}$$

The value represented here again is the marginal cost of accommodating an additional passenger on a train to recoup the costs of the train and its maintenance.

4 Empirical Application

My first application is to estimate the demand curve for German rail travel as detailed above in my theoretical discussion.

4.1 Demand Curve Regression

The majority of the data being used in my analysis was acquired from Eurostat, a Directorate General of the European Union tasked with gathering and providing data on many societal aspects to institutions of the EU and other parties. Data collected by Eurostat was most recently updated in a complete state in 2010. The database contains travel statistics for all cities in the countries of interest and allows the user to pair a departure city with an arrival city and shows the total traffic between the two cities on a given mode of transport in the survey year. From this, I assembled 100 pairs of arrival and departure cities, half with high-speed rail service and half without. These pairs were determined by looking at a map of the German rail network and assembling the pairs, additionally a route was only given a designation of "High Speed" if the route had the high-speed service in 2010 (Raileurope). I then retrieved passenger data for my city pairs on both rail and air travel; I also retrieved density and GDP per capita data for each city, all from Eurostat. Fare price data was also collected for each route for rail transport and air transport. The explicit variables to be used in my regression and their meanings are presented in Table 1 along with their sources. *Totrailcost* and *totaircost* consist of two components, first is the base fare charged to passengers for their journey, then added to this is the opportunity cost which is simply calculated as the time a traveller could be spent working, found by taking the average GDP per capita of the city pair and dividing that by an average annual workload (8 hours a day, 5 days a week). This is then divided by 60 to get the opportunity cost per minute,

Table 1: Regression Variables

Variable	Definition
railpass (Eurostat: Railway Transport)	Total railway passengers between the given city pair
totrailcost (Deutsche Bahn)	The fare cost and opportunity cost of a rail journey in dollars
totaircost (Google)	The fare cost and opportunity cost of an air journey in dollars
density (Eurostat: Density)	The average population density of the given city pair in people/SqKM
gdpcap (Eurostat: GDP per capita)	The average GDP per capita of the given city pair in Dollars
hsr	A dummy variable indicating the presence of HSR service on a given route

and multiplied by the number of minutes in the journey for both air and rail transport.[1] Determining *railpass* merely involved assembling the city pairs in Eurostat and inputting the values into my own dataset. *Density* and *gdpcap* were also straightforward and represent the average of the respective metric in the destination and arrival city. Lastly, *hsr* was determined by seeing which city pairs were positioned on a stretch of high speed line and assigning a dummy value of 1 for high speed service and 0 otherwise[2].

Summary statistics for the variables of interest are presented in Table 2. The first apparent observation that can be made is in regards to costs,

[1]For rail travel time, the total time in transit was merely considered. For air travel time, I added two hours to the total flight time to account for checking in and traveling to airports typically on the outskirts of cities.

[2]In the rare case that a route contained sections both on high speed and normal speed track, the dummy was assigned to the simple majority.

Table 2: Summary statistics

Variable	Mean	Std. Dev.	N
railpass	417255	681150	100
totrailcost	281.03	83	100
totaircost	392.47	134	100
density	2442	630	100
gdpcap	35963	5157	100
hsr	0.49	0.502	100

with air costs at a higher average level than rail costs. Air costs also vary to a greater extent than rail costs with a higher standard deviation at $134 versus $83 for rail. This can be accounted for by looking at some of the lower traveled city pairs, which can still sustain a rail link due to the low marginal costs of rail operation given existing track between the two cities. Realistically, Deutsche Bahn can raise prices on such routes to still keep them affordable, but to offset potential lower demand; air companies cannot do the same and must charge a higher price to recoup costs of flying a low demand route, should the plane not be full. Density and GDP per capita statistics also fall in line with expectations, with GDP per capita being fairly homogenous and density being decidedly less so[3].

I will utilize an OLS multiple linear regression as specified in equation (2) of my theoretical framework. Including the above variables, the regression

[3]This is as a result of looking at the standard deviations and means with standard deviation being 14.3% of the mean for gdpcap and 25.8% of the mean for *density*.

will take the following form.

$$railpass = \beta_0 + \beta_1 totrailcost + \beta_2 totaircost + \beta_3 density + \beta_4 gdpcap + \beta_5 hsr + \epsilon \tag{6}$$

The initial regression indicated a heteroskedastic relationship after applying the Breusch-Pagan Test, thus the following regression summary presented in Table 3 was corrected for heteroskedasticity and is presented with the robust standard errors.

Table 3: Estimation results : Route Demand

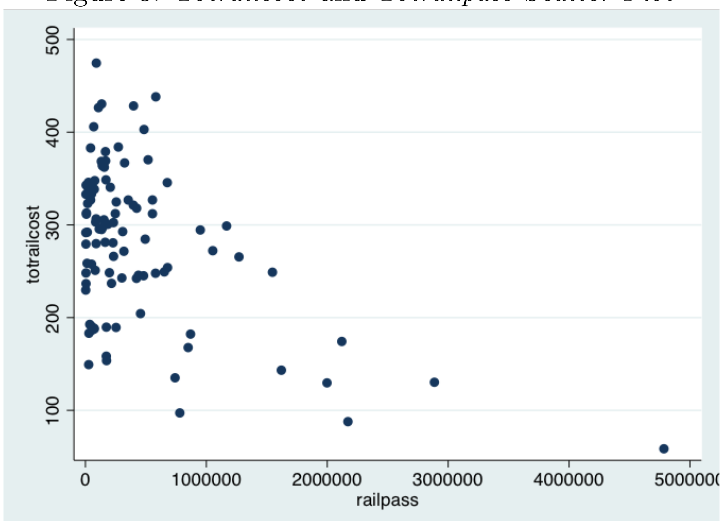
Variable	**Coefficient**	**(Std. Err.)**	**Prob. T**
totrailcost	-3830	(585)	0.000
totaircost	2038	(372)	0.000
density	361	(82)	0.000
gdpcap	12	(10)	0.232
hsr	221484	(103346)	0.035
Intercept	-748015	(445370)	0.096
N	100		
R^2	0.53		
$F_{(5,94)}$	21.212		

The regression results indicate a relationship in line with the expectations of my model, in that a dollar increase in fare price on a railway route will lead to a statistically significant 3830-passenger decrease in annual ridership. Also of inerest is the coefficient on *hsr*, which again is in line with the expectation in that the presence of HSR is estimated to increase annual ridership by 221,484 passengers significant at 98.2% confidence. The variables representing total air cost and density can also be rationally explained. First, the coefficient on *totaircost* is positive and indicates an increase in total air cost of 1 dollar causes substitution of 2038 passengers per year into the competition in that of rail travel, and is significant at 97.2% confidence. Additionally, *density* also shows a positive coefficient of 361 additional annual rail passengers for an additional person/km2 of average density at more than 99% confidence. This can again be rationalized by the notion that the larger average density of the route would imply a greater need for travel among residents. Lastly, the coefficient on *gdpcap* is not found to be statistically significant. An observation of a scatter plot showing the relationship between *gdpcap* and *railpass* indicates a relatively uniform level of rail ridership at all levels of GDP per capita which is intuitively sound when the historical precedent of heavy railway use among Germans is taken into account, particularly among the middle class. This is to say that railway transport is and has been marketed as a mode of transport for all socioeconomic bacgrounds with class based fares, more akin to air travel than public transport. Regardless, the variable will still be included to avoid potential omitted variable bias.

Shortcomings of this portion of the analysis are most evident when looking at the goodness of fit of the linear model estimated. The most straightforward way to check for this is through an application of the Ramsey

RESET test. The test yields an F-value of 20.37, which when compared to an F-statistic of 2.71 at 3 and 91 degrees of freedom indicates that a non-linear relationship has some explanatory power in determining the dependent variable. This is further consistent with a scatter plot of the two main variables of interest, *totrailcost* and *totrailpass* as seen below in Figure 3. A simple observation of the scatter plot shows that there

Figure 3: *Totrailcost* and *Totrailpass Scatter Plot*



is indeed a negative relationship between rail costs and passengers, which seemingly supports the linear estimation given by the regression. However, the rapidly increasing passenger levels correlated with decreasing costs could indicate an exponential or logarithmic relationship between the primary variables of interest. I will however continue with my linear estimation of the demand function with the justification of it being simpler to integrate a linear demand function into my model. Additionally, the RESET test does not preclude a linear relationship but merely indicates that other specifications may be superior, and as such I will continue with the linear model.

Lastly, misspecifications in the model and regression could be caused by problems of multicollinearity, which should not arise due to the nature of my data, but will be looked at anyway. Firstly, the variables do not exhibit strong correlation to each other, and most are again statistically significant. I also calculated the variable inflation factor for the variables, and this yielded values ranging from 1.02 to 1.24 for the variables in the model. These values are interpreted by taking the square root, which then gives the factor by which the standard error of a variable is inflated compared to what it would be if there was no correlation between that variable and others. In the case of my data, these values range from 1.01 to 1.11, and are far below the accepted threshold for high multicollinearity of VIF = 5. This combines to allow me to reasonably conclude that multicollinearity of my variables is not an issue in my regression.

4.1.1 Demand Curve Estimation

I now specify the explicit demand curve according to the method laid out in Section 3.2 and equation (3).[4] Again, the regression I ran above estimates the effect of total costs on ridership levels, and would place cost as independent and ridership as dependent if merely placed into a demand function, essentially showing the inverse demand function. I could opt for

[4] Again, with the values from the regression being transformed according to

$$TC = \frac{\beta_0}{\beta_1} + \frac{1}{\beta_1} D \tag{7}$$

regressing all the variables on price instead to bypass this step but the coefficients would be different and not represent the relationship that I am ultimately after. Additionally, my model utilizes the other control variables to determine a more precise estimate for the coefficient on *railpass* and the alternate model would not indicate this effect in my desired relationship.

To continue, according to equation (2) my demand function for German railway routes is,

$$TC = 195 - \frac{1}{3830}n \tag{8}$$

with TC representing the total cost of rail usage, and n representing the passenger level. This simply indicates that starting at a peak price of \$195 (vertical intercept), a one-dollar decrease in price will increase ridership by 3830, to a horizontal intercept of 746,850 passengers. The inclusion of the effects of *HSR* into this will shift this demand curve upward to a new vertical intercept of 253, while still retaining the same slope and pushing the horizontal intercept out to 968,990. This comes to indicate that the presence of HSR on a given route will increase demand for rail travel on that route by approximately 23%.

This is greater than the estimated increase in demand of 9% referenced by Cuoto and Graham earlier, but we must remember that this level was estimated for the impact of HSR on national railway demand rather than route level demand. This however is not out of line with intuition, as a national presence of HSR might be small relative to the size of a network. Thus only a small portion of the population that lives along the routes that

HSR services and has access to it and their increased demand would be counteracted by the relatively stable demand on other routes with no HSR service. Conversely, my higher estimate for route level demand increase reflects the (reasonably) larger increase in demand of people who have direct access to the good. This demand function will be revisited later when the full model is assembled, and the following section will focus on determining the cost curves in my model.

From this demand curve, I can also simply determine the marginal revenue curve necessary for the model through basic economic relationships. Given that total revenue = price x quantity,[5] I can take my demand function and multiply by quantity as follows to determine total revenue, multiply quantity through the equation, and take the derivative to yield marginal revenue as follows.

$$TR = TC = (195 - \frac{1}{3830}n)n \ \ or \ TR = 195n - \frac{1}{3830}n^2 \tag{9}$$

$$MR = \frac{\delta TR}{\delta n} = 195 - \frac{1}{1915}n \tag{10}$$

This marginal revenue curve maintains the same intercept as the total cost but with half of the slope and represents the marginal revenue of each additional passenger on a given route. It will be revisited and incorporated into the overall model to determine the monopolist price level, and the required subsidy level.

[5]With my variables of TC and n equaling price and quantity respectively.

4.2 Cost Curve Estimation

The estimation of cost curves for the model is generally fairly straightforward, but still require some explanation. With demand estimated, I will first focus on average fixed costs. In my case, AFC is being determined purely as the infrastructure construction costs necessary to build a given routes track at a given passenger level. This can also be equated to overall average costs, as I noted in Section 3 with my assumption of negligible variable costs. To reiterate, Germany is an ideal choice on these grounds for this analysis as all German trains have the ability to run on all types of track, unique to the German rail network and allowing me to utilize a simple average construction cost per mile of track to apply to my entire analysis. First, to determine a total variable for route cost I took the average cost per mile in dollars of the network (23.1 million per mile) and multiplied this cost by the length of each route in miles to yield a variable for total route cost (Feigenbaum, 2013). These total route costs are then divided by 10 years, which is the the amortization period of loans used to fund the construction and the average is taken, giving an average route cost per year of \$54.18 million.

To determine the actual formula for AC I will follow the method specified in section 3.3 and following equation (4). This simply takes the above value of average cost per year and divides it by passenger level (n) to determine the average costs at a given passenger level. The explicit form taken by the AC curve is as follows.

$$AC = \frac{54180000}{n} \tag{11}$$

This gives a downward sloped and convex AFC curve, which as discussed previously, can be equated to ATC, given my assumption of negligible variable costs. The curve itself determines the necessary subsidy level, with the monopolist passenger quantity being taken up to the AFC curve, which gives the price level necessary to charge riders to recoup infrastructure costs.

Marginal costs are held constant in my model and are represented as the cost of accommodating an additional passenger on any given route covering the costs of rolling stock and maintenance. This fixed level of marginal cost per route passenger will be determined according to equation (5) that again takes the total cost of all rolling stock and divides it by the number of years over which the funds used to purchase the trains are amortized. Then the cost per year is divided by the maximum theoretical ridership to determine the marginal cost of accommodating an additional passenger. The calculation is very straightforward, but values and intuition of results are still necessary to discuss. Values for the cost of train units are determined from Siemens AG, the provider of all rolling stock. Converted into dollars, the approximate cost per train unit purchased by DB is $41,500,000 and is paid for over 10 years with maintenance included in the purchase price (Siemens, 2008) Additionally necessary to account for are energy and labor costs per train which are estimated to be $1,168,000 per train in energy and $800,000 per train in labor costs which are added to the cost per unit (Feigenbaum, 2013). This is then multiplied by 67, or the number of recently (since 1998) purchased train units in operation to yield a total cost of rolling stock and maintenance of approximately $2,912,356,000 which when divided by 10 years gives an annual cost of $291,235,600.

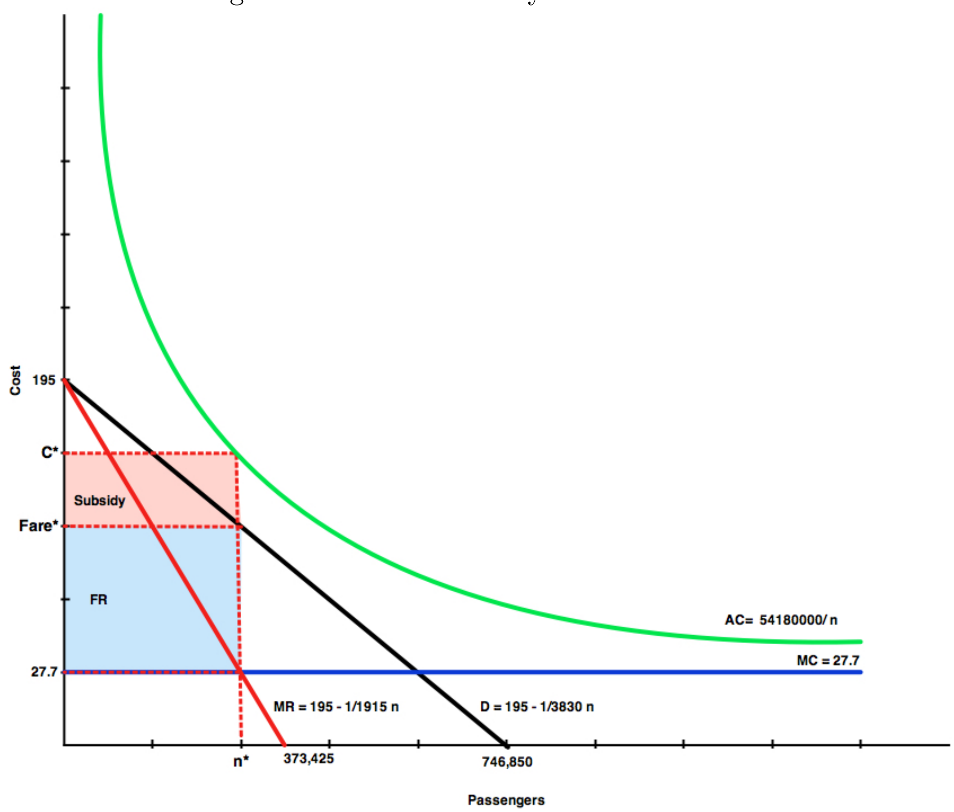
The total annual cost must then be divided by maximum theoretical annual ridership to determine the marginal cost of an additional passenger. The maximum capacity of one DB train unit is on average 430 passengers, and in between trains in maintenance time, and rotation of older train stock on routes, I will assume each new train makes an average of one journey per day. The average capacity per train of 430 is multiplied by 67 trains and 365 days for a total of 10,515,650 possible passengers in a given year. When the annual cost is divided by the maximum possible ridership it is determined that the marginal cost of an additional passenger is $27.70. Intuitively, this indicates that for any additional passenger on any route, $27.70 of their fare price will always be required to cover the train units and their operation. The next section will focus on incorporating this marginal cost curve and all previous curves into my model and determine the change in subsidy level resultant from a shift in the demand curve brought about by the presence of HSR.

5 Cost-Benefit Analysis

5.1 Subsidy Calculation

Now, having derived all the necessary curves, I can assemble the following graph for my model, representing a given German railway route without HSR service, showing demand and costs. To note, the graph is not drawn to scale due to the high passenger levels relative to costs. The main purpose of this section is to look at the subsidy levels required to make a given route financially viable by offsetting infrastructure costs and supplementing

Figure 4: German Railway Route- No HSR

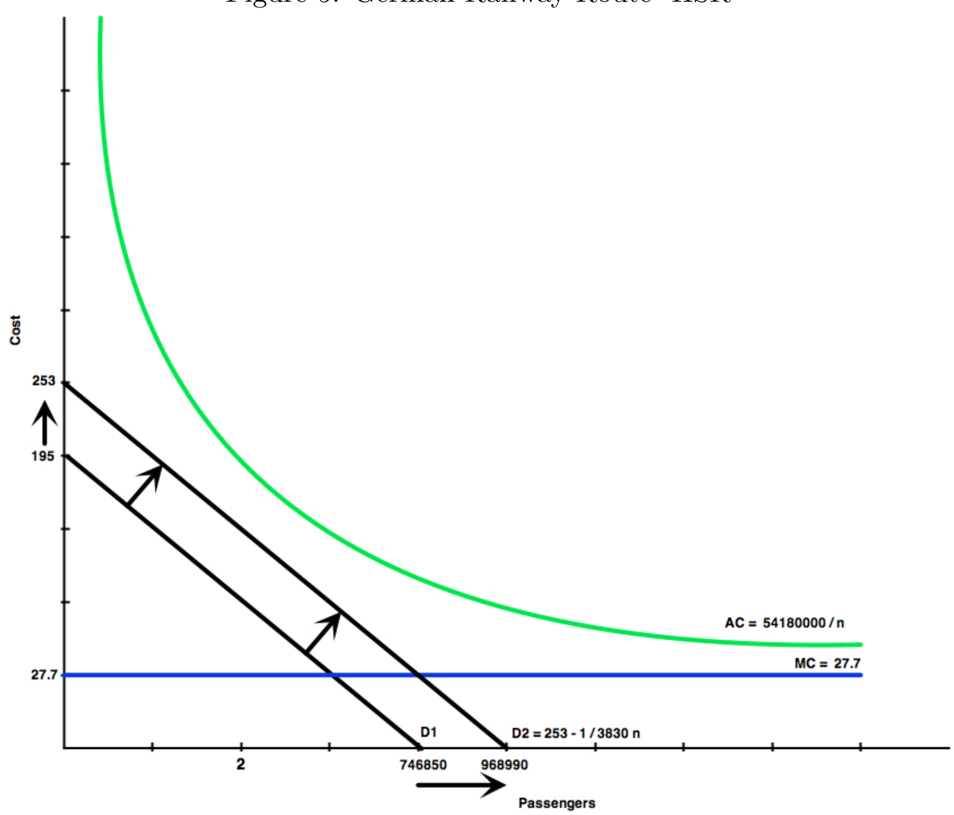


fare revenues. In Figure 4, this is represented as the shaded area marked subsidy, comprising of the difference between fare revenues (FR) at the monopolist price and the average cost curve. To begin the calculation, I will first determine the equilibrium level of passengers n*, on a given route by setting marginal revenue equal to the marginal cost of 27.7 and solving for n, which in this case yields a value of 320,380 passengers. This n* is then inputted into the demand and AC equations to give the equilibrium fare and infrastructure cost not covered by fare revenue which are \$111.35 (Fare*) and \$169.11 (C*), respectively. Next, I will determine fare revenue by taking the difference of Fare* and marginal costs to give a revenue per passenger of 83.65, this is then multiplied by the equilibrium ridership level n^* to give a fare revenue of approximately \$26.8 million per annum on a given route without HSR, represented in Figure 4 by the shaded region marked FR. The required subsidy level is calculated similarly by taking the difference of C* and Fare* and then multiplying by n^* to get a reqiured annual subsidy of \$18.505 million per annum on a route.

The main statistic that this boils down to is that for a given non high speed route in Germany, an annual subsidy of 34.2% of total costs is required to supplement fare revenues and break even at the equilibrium price and ridership level. In hindsight of looking at several other studies of current required transit subsidies, my estimate would seem to understate the actual level of subsidization present in much of Europe and America. For example, rail projects in the UK and France typically subsidize total costs by two thirds to arrive at the final passenger fare (European Environment Agency, 2007). In America the levels are even higher, with an average of 70-80% of costs being subsidized on commuter rail projects (Garrett, 2004). This can lead to two conclusions, first that my figures and model are miss-specified, which may account for some of the variation due to previously mentioned simplifications I make, such as assuming demand to be linear and variable costs to be negligible. Second, the nature of these cost-benefit analyses in general may result in an understatement of the required subsidy levels (Utsunomiya and Hodota, 2011). This is to say that all cost-benefit valuations of existing systems undertaken to make future policy suggestions are chronically unable to account for the full costs of rail development in terms of cost overruns, corruption, and inefficient allocation of funds to truly come to a reasonable policy conclusion based on ex-post data.

Still, the main goal of this study is to estimate the effect of HSR on the reqiured subsidy level according to my model. As noted previously in the discussion of equation (10), the inclusion of the effects of HSR service on a given route increase demand by about 23% through the addition of the *hsr* variable to the demand function. This increases the vertical intercept to 253 and retains the original slope for a new horizontal intercept of 968,990 passengers, the shift can be seen in figure 5.

Figure 5: German Railway Route- HSR



This also shifts the marginal revenue curve as it is derived from demand up to a vertical intercept of 253 and again at half the slope of demand, then the same calcualtions as above are applied to determine fare revenue and subsidy level.When equating marginal cost with the new marginal revenue curve, a new passenger level of 431,450 is given. This is then put into the new demand curve and the original cost curve to yield a fare and cost per passenger of $112.65 and $125.58 respectively. Then fare revenue and subsidy is calculated in the same manner as before to yield a fare revenue of approximately $36,651,700 and a required subsidy of $5,578,650. This indicates a required annual subsidy that is 10.3% of total annual costs. The key statistic that I have reached in this section is the difference in required subsidies between the HSR and non-HSR routes, which is found to be 23.9%, thus indicating the presence of HSR (through its effect on demand) causes a 23.9% decrease in required rail subsidy. This is intuitively sound and in line with my assumed model from Section 3, allowing me to conclude that the rough model can indeed be applied to the German railway market as a representation of its basic operation at the route level.

5.2 Model Shortcomings

There are some potential places for improvement and possibilities for shortcomings in the model however. Most significantly, the magnitudes of shifts and changes in demand and subsidies are indeed fairly robust, but are not very specific nor precise. The biggest shortcoming is most likely in my estimation of average costs, which are understating their actual levels as discussed above. There are indeed variable costs associated with rail travel, but I assumed them away for simplicity of the model, and with the rationalization that they are relatively small compared to total infrastructure costs, but nonetheless they are there and would increase average cost per passenger. Additionally, my level of total route cost is also understating actual total costs. Realistically, total cost would have to be measured for each individual route and then regressed to determine a more fitting estimate for these costs for a given route. This is because I took a mere average construction cost per mile and then averaging this for each route, in reality the route construction costs vary wildly in regards to environmental and urban obstructions that must be negotiated on some routes, in rare cases being as much as double the average.

Another factor to consider is cost overruns and other expenses that are not reported in some cases as part of the total construction costs, all of which combine to indicate that my explicit example understates the total costs of infrastructure, and in turn average costs per passenger. This indicates that a more realistic account of costs would then lead to larger required subsidies to make a route financially viable, combined with my estimated demand. On the note of demand, in preliminary observation, a logarithmic relationship would be better suited to my dataset and yields a closer fit, but I proceeded with the linear regression to maintain model simplicity. Along these lines, due to the convexity of the logarithmic function and an observation of the scatter plot of the data points, I would conclude that I slightly overstated the central range of passenger demand.

All of this combines to indicate what was mentioned previously, in that most of these studies do not accurately represent the full costs and required subsidies necessary to support a rail system. However, it is still reassuring to see that the directions of changes in the model all move as expected. As far as policy implications, the true decisions lie with policy makers being able to rationalize the cost to taxpayers against revenues brought in through fares. The goal of policy makers in this regard is to ensure that demand and fare revenues are maximized while costs and required subsidies are minimized, to pass on a minimum of cost to taxpayers and lessen the double payment made by riders, who essentially pay twice for their fares and in taxes. This conclusion is not reached in my study, but with a more accurate representation of costs, the general model could be used to determine this level and comment on the effectiveness of the implementation of the system in Germany.

6 Conclusion

Much of the work done to analyze the costs and benefits of HSR to comment on policy choices presents an overly optimistic view and in many cases fails to account for the full costs of a rail project. Given a natural monopoly market structure with fixed costs that exeed revenues for all levels of demand, it is difficult to comment on the viability of HSR investment without context as to the financial standing of a country and the ability of policy makers to justify the cost to taxpayers. Looking at Germany, I worked to apply a general model of railway subsidization to an ideal representation of a high speed rail network with readily available data.

In my case, I found that Germany does conform to my model and additionally to previous studies such as that by Cuoto and Graham and exhibits an increase in consumer demand on a given route with the presence of HSR service. Required subsidies also decrease through a result of my specified model, but there is much room for improvement upon my basic framework. Most significantly, an accurate account of all costs combined with a better specification of costs in the model would yield levels that are more in line with real world observations of costs that require subsidies of over 50%.

Assuming that my representation of a given German route is accurate, I could reasonably reccommend for further high speed development as opposed to basic railway development and maintenance due to a decrease in subsidies and increase in fare revenues brought about through increased demand. But again, before other developments proceed, such as that in California, I would more conservatively suggest for analysts to look closely at full costs and any potetnial for cost overruns to not understate cost levels, and also to make sure an accurate representation of consumer demand is made to not overstate demand. This will ultimately give policy makers a more accurate representation of required subsidies and allow them to better justify these costs to taxpayers.

References

[1] Albalate, Daniel and Germa Bel. "High Speed Rail: Lessons for Policy Makers from Experiences Abroad." *Public Administration Review*, 72 (2012): 336-349. Web.

[2] Behrens, Christiaan, and Eric Pels. "Intermodal Competition in the LondonParis Passenger Market: High-Speed Rail and Air Transport." *Journal of Urban Economics*, 71.3 (2012): 278-88. Web.

[3] Couto, Antonio, and Daniel J. Graham. "The Impact of High-speed Technology on Railway Demand." *Transportation*, 35.1 (2007): 111-28. Web.

[4] "Deutsche Bahn at a Glance." *Deutsche Bahn*. Web. 28 Oct. 2014. http://www.bahn.de/i/view/GBR/en/about/overview/company_profile.shtml.

[5] Domencich, Thomas A., and Gerald Kraft. *Free Transit*. Lexington, MA: Heath Lexington, 1970. Print.

[6] Dunn, James A., and Anthony Perl. "Policy Networks and Industrial Revitalization: High Speed Rail Initiatives in France and Germany." *Journal of Public Policy*, 14.03 (1994): 311. Web.

[7] Eurostat: Density [demo_r_d2ens]. "Population density NUTS 3 regions." *Eurostat*, (2012). Web. Database code demo_r_d3dens.

[8] Eurostat: GDP [nama_r_e2gdp]. "Gross domestic product at current market prices by NUTS 2 regions." *Eurostat*, (2011). Web. Database code - nama_r_e2gdp.

[9] Eurostat: Railway Transport [tran_r_rapa]. "National and international railway passengers transport by loading/unloading NUTS 2 region." *Eurostat*, (2010). Web. Database code tran_r_rapa.

[10] Feigenbaum, Baruch. "High-Speed Rail in Europeand Asia: Lessons for the United States."*Reason Foundation*. Policy Study 418, May 2013. Web.

[11] Garrett, Thomas. "The Costs and Benefits of Light Rail." *The Central Banker*, Fall (2004). Web.

[12] "Flights - Google Search." *Flights - Google Search*. N.p., n.d. Web. 28 Oct. 2014. https://www.google.com/flights/.

[13] Kuzmyak, J. Richard, and Eric N. Schreffler. *Evaluation of Travel Demand Management Measures to Relieve Congestion.* U.S. Federal Highway Administration Report No. FHWA-SA-90-005. Springfield, VA: National Information Service. 1990. Print.

[14] Levinson, David, Jean Michel Mathieu, David Gillen, and Adib Kanafani. "The Full Cost of High-speed Rail: An Engineering Approach." *The Annals of Regional Science*, 31.2 (1997): 189-215. Web.

[15] Mohring, Herbert. "Optimization and Scale Economies in Urban Bus Transportation." *The American Economic Review*, 62.4 (1972): 591-604. Web.

[16] "Germany : DB." Rail Europe. *Rail Europe,* n.d. Web. 28 Oct. 2014. http://www.raileurope-world.com/about-us/railways/article/germany-db.

[17] Rus, Gins De, and Vicente Inglada. "Cost-benefit Analysis of the High-speed Train in Spain." *The Annals of Regional Science*, 31.2 (1997): 175-88. Web.

[18] Savage, Ian. "A Comment on 'Subsidisation of Urban Public Transport and the Mohring Effect." *Journal of Transport Economics and Policy*, 44.3 (2010): 373-80. Web.

[19] "Siemens Receives Order over 15 High-speed Trains from Deutsche Bahn." *Siemens*. 17 Dec. 2008. Print.

[20] Siemens AG. Communications and Government Affairs. *ICE Trains for Deutsche Bahn Approved for Germany.* Munich: Siemens, 2013. Print.

[21] Siemens, Deutsche Bahn, and the Federal Ministry of Transportation. *Deutsche Bahn and Siemens Present New ICE 3 in Berlin.* Berlin: Siemens, DB, and BMVI, 2014. Print.

[22] Small, Kenneth A. *Urban Transportation Economics.* Chur: Harwood Academic, 1992. Print.

[23] United States. Federation of American Scientists. Congressional Research Service. *High Speed Rail in the United States.* By David R. Peterman, John Frittelli, and William J. Mallett. Congressional Research Service, 08 Dec. 2009. Web.

[24] Utsunomiya, Mariko, and Kenichi Hodota. "Financial Lessons from Asian Experience in Constructing and Operating High Speed Train Networks." *Transportation*, 38.5 (2011): 753-64. Web.