CEPR Centre for Economic Policy Research [redacted] The Pinhas Sapir Center for Development Tel Aviv University [redacted] TEL AVIV UNIVERSITY The Gordon Faculty of Social Sciences Department of Public Policy

EIGHTH CEPR/JIE SCHOOL ON APPLIED INDUSTRIAL ORGANIZATION

Hosted by

The Eitan Berglas School of Economics, Recanati School of Business, Department of Public Policy at Tel Aviv University

Supported by

The Pinhas Sapir Center for Development, Tel Aviv University
Recanati School of Business, Tel Aviv University
Department of Public Policy, Tel Aviv University
Journal of Industrial Economics (JIE)
CEPR

Tel Aviv; 24 May 2011

Demand Estimation with Selection Bias due to Exit/Takeovers: A Dynamic Approach with an Application to the U.S. Railroad Industry

Daniel Coublucq (Toulouse School of Economics)

Demand Estimation with Selection Bias due to Exit/Takeovers: A Dynamic Approach with an Application to the US Railroad Industry[1]

Daniel COUBLUCQ[2]

Toulouse School of Economics

May 15 2011

ABSTRACT

This paper is motivated by the US freight railroad industry which is characterized by a major restructuring over the last 30 years In particular the number of active firms decreased from 26 in 1978 to seven in 2006 due to several takeover waves Our empirical focus concerns the estimation of a structural demand model for the US railroad industry and then we use these estimates to compute the evolution of the mark ups the quality of the freight services provided and the consumer surplus The restructuring of this industry involves significant exit and takeovers This implies that our data is characterized by an attrition issue which generates a selection problem A focus is to provide an estimation algorithm which takes explicitly into account this attrition issue We find that our algorithm produces more plausible estimates of demand coefficients compared to standard estimation procedures Moreover using our model we recover the evolution of marginal costs mark ups and consumer surplus over time We find that the takeover waves have led to efficiency gains by decreasing the marginal costs and this was translated into lower prices and an increase in the consumer surplus Finally the takeovers have led to a reallocation of assets from the less efficient firms to the most efficient firms which improved the quality of the freight services provided

Keywords: selection bias panel data demand model merger takeover analysis railroad industry
JEL Classification: C23 C51 L10 L41 L92

---

[1] I would like to thank my advisor Marc Ivaldi for his encouragement and suggestions during the process of this work I would also like to thank Gerard McCullough for providing the data and his guidance regarding the data construction Finally I would like to thank Mehtap Akgü Sofronis Clerides Peter Davis Yinghua He Thierry Magnac the participants at the ZEW conference Quantitative Assessments in Competition Policy " at the DIW conference End of Year Summit 2010 " and at the ENTER Jamboree 2011" in Tilburg University for useful comments All remaining errors are mine A previous version of this paper has circulated under the name Demand Model and Selection Bias due to Exit Mergers: A Dynamic Approach with an Application to the U S Railroad Industry "

[2] E mail: daniel coublucq tse fr eu Toulouse School of Economics 21 All e de Brienne 31000 Toulouse France

1 Introduction

The US railroad industry is characterized by a rather “light” regulation This regulatory freedom came from the Staggers Act which deregulated the US railroads in 1980 In particular this deregulation process came with several takeover waves that have led to a concentrated industry today As Figure 1 and Figure 2 illustrate there were 26 firms in 1978 while there are seven firms today see also Table 1and Table 2)

The economic motivation of this paper is to examine the impact of past takeovers on the consumer welfare In this framework the consumers could be the farmers the manufacturers and so forth who ship the goods to another place For example how has the concentration process impacted the consumer welfare over time? As mentioned in Whinston 2007) a retrospective analysis is a critical step for understanding past takeovers and improving future enforcement practice This can be related to a current debate regarding the state of competition in the US railroad industry see the GAO Report 2006)) In this vein this paper analyzes the consumer welfare and its evolution by taking into account the past takeover waves hence the concentration process between 1980 and 2006 in the US railroad industry While doing this three important dimensions of consumer welfare is highlighted

The first one is related to the pricing of the railroad companies and its implications for the consumers According to the trade off suggested by Williamson 1968) if the consumers have benefited from this concentration process one might conclude that the railroad firms have passed the productivity gains due to takeovers to the consumers through a decrease in the final prices Otherwise it might be that the railroad firms have used the concentration mainly to decrease the competitive pressure and increase their mark ups

Secondly although the main variables of interest are the price of freight services and its evolution over time with its impact on the consumer surplus it is also important to incorporate the investment of the firms in the network since it impacts the quality of the network infrastructure and might be altered by the takeover waves To this end we construct a variable that represents the quality

of the network infrastructure which impacts the quality of the freight services provided and hence enters into the consumer welfare function

The last but not the least aspect of the paper during the analysis of the consumer welfare is to consider the takeovers as a way to reallocate assets from firms with low asset exploitation abilities to firms with high exploitation abilities In this sense our dataset on the US railroad industry is particularly suited for this purpose since this industry is characterized by an important attrition due to exit and takeovers from 26 firms in 1978 only seven firms remain in 2006) Moreover this redeployment of assets might lead to a selection bias in the sample since the least efficient firms leave the industry and the more efficient ones stay[3] This is an essential issue that this paper takes into account

In particular we provide an estimation algorithm to take into account this selection issue using a dynamic model which allows for the equilibrium force of exit and from which a structural selection equation is derived The selection equation yields a correction term which is then plugged into the structural demand model in order to correct for selection bias Intuitively this can be seen like a standard two step Heckman procedure where the correction term is derived from an explicit dynamic model of exit behavior While doing this we build on the dynamic models of Aguirregabiria and Mira 2007) Bajari Benckard and Levin 2007) Berry Ostrovsky and Pakes 2007) Dorazelski and Satterthwaite 2010) Ericson and Pakes 1995) and on the literature on selection bias Wooldridge 1995) Semykina and Wooldridge 2005)) The idea of using a structural model to account for selection is also an important contribution of Olley and Pakes 1996) in the framework of production function Another related paper is Mazzeo 2002) where a structural model of entry is used to deal with endogeneity of the market structure variables i e the number of product type competitors in a reduced form price regression

[3] The interested reader can look at Andrade et al 2001) Andrade and Stafford 2004) Jovanovic and Rousseau 2002) Maksimovic and Phillips 2001) and Olley and Pakes 1996) This literature shows that mergers takeovers are a way to reallocate assets and achieve productivity gains since the least efficient firms leave the industry and sell their assets to the most efficient firms and thereby put the resources to their best use

Furthermore we also emphasize the connection between the demand model and the dynamic exit model In particular the variables that enter the demand model influence the spot profit function and thus the value function In other words all the variables that are in the demand model must also appear in the dynamic model To deal with this issue we carefully define the observed and unobserved state variables in order to ensure that the demand model is fully coherent with the dynamic model and this leads to an iterative algorithm Lastly we also explicitly discuss the link between the exogenous process of the unobserved state variable and selection bias Indeed the unobserved state variable plays an important role and it is interpreted as the unobserved efficiency ability of a firm to provide freight services of good quality

Moreover using the estimates of the parameters of the demand model we are able to recover the evolution of the marginal costs mark ups and consumer surplus over time We find that the takeovers led to important efficiencies with an important decrease of the marginal costs during the period 1980 2004 This decrease in the marginal costs was translated into a decrease in the freight prices for the same period For the period 2004 2006 which is characterized by a price increase we argue that it might be due to an increase in the industry marginal cost This implies that the increase in freight price was rather a consequence of an increase in the marginal costs instead of an abuse of market power by the railroad firms This is relevant for the current debate regarding the potential market power of the railroad firms in the US see GAO report 2006)

Furthermore during the takeover waves the consumer welfare has increased which can be justified by two elements The first element is the cost efficiency gains Indeed this industry is characterized by important returns to density see Berndt et al 1992 1993a 1993b) Ivaldi and McCullough 2001 2008)) In addition to the gains from the consolidation of the traffic Dennis et al 2010) mention other efficiency gains of takeovers in the sense of more efficient use of equipment by rerouting traffic to avoid congestion on routes reduction of ton miles by choosing the shortest routes routes dedicated to specialized services and so forth This led to a decrease in marginal costs which resulted in a decrease in final prices The second element is the efficiency gains in terms of the

provision of the freight services which is the ability of firms to provide freight services of good quality due to the reallocation of assets from the

takeover waves that led to a concentrated industry today Namely there were 26 firms in 1980 while there are only seven firms today see Waters 2007))[4]

Railroads class 1 and regional railroads) account for 41% of freight ton miles more than any other mode of transportation Figure 4 illustrates the increasing importance of Class 1 railroads in the US freight market Coal is the most important commodity carried by US railroads In 2007 coal accounted for 44% of rail tonnage and 21% of rail revenue Coal accounts for around half of the US electricity generation and railroads handle more than two third of US coal shipments Other major commodities carried by rail include chemicals ethanol plastic resins fertilizers) agricultural products such as grain non metallic minerals such as phosphate rock sand and crushed stone food products steel forest products e g paper) motor vehicles and motor vehicle parts waste and scrap materials

Several papers have studied the US railroad industry from a cost side perspective The cost function analyses of the 1990s —Friedlaender 1992) Berndt et al 1993a) Berndt et al 1993b)— show that the economies of density are significant and substantially greater than one[5] This finding is confirmed by Ivaldi and McCullough 2001 2008) in a different framework in which they consider three different types of freight services namely bulk intermodal and general freights measured in car miles instead of ton miles They find economies of density equal to 1 90 This seems to be a pervasive technological characteristic of the US railroad industry which yields a natural tendency for firms to be concentrated

Another interesting feature of this industry is the capital adjustment since the deregulation in 1980 Berndt et al 1993b)) Indeed the Staggers Act provided the railroads with considerable

---

[4] Namely they are: Burlington Northern and Santa Fe Railway Company BNSF) Kansas City Southern Railway Company KCS) Union Pacific Railroad UP) Soo Line Railroad Company SOO) which represents the U S operations of the "Canadian Pacific" railways company CSX Transportation Inc CSX) Norfolk Southern Combined Railroad Subsidiaries NS) Grand Trunk Corporation GTC) which represents the U S operations of the "Canadian National" railways company Source: Surface Transportation Board STB)

[5] The returns to density are defined as $S_y = \left(\partial \ln VC / \partial \ln y\right)^{-1}$ where VC represents the variable cost and y represents the output in ton miles In the paper mentioned the returns to density in terms of mean value for each firm) have a range from a minimum of 1 477 to a maximum of 4 274

potential to rationalize their capital structure [6] Specifically Berndt et al 1993b) look at the extent of capital adjustment after the Staggers Act They find that cost savings from increments in ways and structures capital are not enough to justify the observed levels of way and structures capital This implies an excessive capacity in the industry As mentioned by Berndt et al 1993b) as the regulatory freedom allows the firms to adjust their capital the overcapitalization of the railroad firms seems puzzling Berndt et al 1993a) mention the possibility that cost savings from increments in way and structures capital may not fully reflect the benefit of the investment In particular it is the case if service quality enters the demand function and if it depends on the amount of way and structures capital Then the shadow value of capital which considers only a cost perspective) underestimates the actual benefits of investment In addition to reducing costs the way and structures investment may enhance demand by allowing higher speed and better service thanks to the high quality rail This can be used to justify the inclusion of network quality in the demand model see section 4 regarding the demand model and section 5 for the data construction)

3 A dynamic model to correct for selection bias

As mentioned in Olley and Pakes 1996) and in Wooldridge 2002) when the panel is unbalanced it is important to account for the nonrandom nature of the sample and this requires a formal description of why the panel may be unbalanced The purpose of this section is to use a dynamic model of firm behavior which allows for the equilibrium force of exit in order to derive a correction term that will allow estimating consistently the parameters of the demand model

We consider a model of dynamic competition between oligopolistic competitors We build on Ericson and Pakes 1995) Berry Ostrovsky and Pakes 2007) and Dorazelski and Satterthwaite 2010) The key feature is that actions taken in a given period may affect both current profits and by influencing a set of commonly observed state variables future strategic interactions Thus this model

6

permits the aspects of dynamic competition such as exit and investment decisions As mentioned in the introduction we focus on the exit decision since it allows us to derive a correction term to deal with selection bias in the estimation of the demand model

We consider infinitely lived agents that make decisions at times $t = 1 \ldots +\infty$ where J_t denotes the number of agents at time t Conditions at time t are summarized by a vector of state variables $\boldsymbol{w}_t \in W = \mathbb{R}^{2J_t+1}$ where $\boldsymbol{w}_t = J_t \; \nu_{1\,t} \ldots \nu_{j\,t} \ldots \nu_{J_t\,t} \; \xi_{1\,t} \ldots \xi_{j\,t} \ldots \xi_{J_t\,t}$ where $\nu_{j\,t}$ represents the observable quality and $\xi_{j\,t}$ represents the unobservable quality or efficiency of freight services provided by firm j at the beginning of period t Given the state $\boldsymbol{w}_t$ firms choose their actions simultaneously These actions include three components: a pricing decision an investment decision and an exit decision In the model the pricing decision is static whereas investment and exit are dynamic decisions

At the beginning of each period each incumbent firm is assigned a random scrap value received upon exit Scrap values are privately known that is whereas a firm learns its own scrap value prior to making its decisions its rivals' scrap values remain unknown to it Adding firm heterogeneity in the

the state variables and will be interesting for the estimation of the investment policy function However our paper deals with selection bias due to exit and as we show below only the second type of heterogeneity in the scrap value will be interesting for our purpose Thus we do not enter into the details regarding the evolution of the state variables Second we introduce heterogeneity in firms by taking into account the differences in opportunity costs of staying in the industry Indeed we assume that incumbents have random scrap values that are independently and identically distributed across firms and over time

The within period timing is the following:

- At the beginning of date t each incumbent firm learns its scrap value denoted $\phi_{j\,t}$ and decides on exit and investment
- Firms compete in the product market by making pricing decisions based on the vector of state variables $\boldsymbol{w}_t \equiv J_t\ v_{1\,t}\ \dots\ v_{j\,t}\ \dots\ v_{J_t\,t}\ \xi_{1\,t}\ \dots\ \xi_{j\,t}\ \dots\ \xi_{J_t\,t}$
- At the end of period t the vector of state variables is updated and firms take their decisions in period $t+1$ conditional on $\boldsymbol{w}_{t+1}$ Regarding the exit decisions the number of active firms at period $t+1$ is equal to J_t minus the number of firms that exit at period t In addition if the firm j does not exit the market at period t the investment decision is carried out and its state variables are updated

We assume that at the beginning of the period each incumbent receives a random scrap value $\phi_{j\,t}$ from a distribution F with $E\ \phi_{j\,t} = \phi$ and the scrap values are identically and independently distributed across firms and over time Incumbent firm j at time t learns its scrap value prior to its exit and investment decisions but the scrap value of its rivals remain unknown to it Let $\chi_{j\,t}\ \boldsymbol{w}_t\ \phi_{j\,t} = 1$ indicates that the incumbent firms chooses to stay in the industry in state $\boldsymbol{w}_t$ and $\chi_{j\,t}\ \boldsymbol{w}_t\ \phi_{j\,t} = 0$ denotes the firm j chooses to exit the industry at time t Because this decision is conditioned on its private scrap value it is a random variable from the perspective of the other firms

We use $\varsigma_{j,t}(w_t) = \int \chi_{j,t}(w_t, \phi_{j,t}) dF(\phi_{j,t})$ to denote the probability that firm j at date t remains active in the industry in state w_t

We assume that the firms compete in prices in the product market The two state variables of a firm j at time t are denoted $w_{j,t} \equiv (v_{j,t}, \xi_{j,t})$ The optimal price is a function of the quality level of the firm j ($w_{j,t}$) and the quality levels of its competitors $w_{-j,t} = (w_{1,t}, \ldots, w_{j-1,t}, w_{j+1,t}, \ldots, w_{J_t,t})$ Thus the profit function from the product market competition for firm j at time t is a function of the industry state $w_t \equiv (J_t, w_{j,t}, w_{-j,t})$, $\pi_{j,t}(w_t)$ We will come back on this point in the next section when we present the demand model

In addition to receiving profit the active firm incurs the cost of investment $c(i_{j,t})$ where the investment level $i_{j,t}$ is decided at the beginning of the period The active firm moves then to the next period state according to the transitions of the state variables

The probability that the industry transits from state w_t to state w_{t+1} depends on the exit and investment decisions of the firms at date t More formally we consider the transition function $P(w_{t+1} | w_t, \boldsymbol{\chi}_t(w_t, \boldsymbol{\phi}), \boldsymbol{i}_t(w_t))$ where $P(\cdot)$ is the transition probability that the industry moves from state w_t to state w_{t+1} given the firms' decisions to stay or to exit that is $\boldsymbol{\chi}_t(w_t, \boldsymbol{\phi}) = (\chi_{1,t}(w_t, \phi_{1,t}), \ldots, \chi_{J_t,t}(w_t, \phi_{J_t,t}))$ and given the investment decisions that is $\boldsymbol{i}_t(w_t) = (i_{1,t}(w_t), \ldots, i_{J_t,t}(w_t))$ As mentioned in Dorazelski and Satterthwaite (2010), the exit decision is a random variable since a firm's scrap value is private information thus we integrate out over all possible realizations of firms' exit decisions to obtain the probability that the industry transits from state w_t to state w_{t+1}:

$$\int \cdots \int P(w_{t+1} | w_t, \boldsymbol{\chi}_t(w_t, \boldsymbol{\phi}), \boldsymbol{i}_t(w_t)) \prod_{j=1,\ldots,J_t} dF(\phi_{j,t}) \qquad (1)$$

Since the scrap values are independently distributed across firms, a particular realization of firms' exit decisions, denoted $\boldsymbol{l}_t = (l_{1,t}, \ldots, l_{J_t,t})$ where $l_{j,t} = 1$ if the firm j stays and $l_{j,t} = 0$ if firm j exits, occurs with probability $\prod_{j=1,\ldots,J_t} \varsigma_{j,t}(\boldsymbol{w}_t)^{l_{j,t}} (1-\varsigma_{j,t}(\boldsymbol{w}_t))^{1-l_{j,t}}$ where $\varsigma_{j,t}(\boldsymbol{w}_t) = \int \chi_{j,t}(\boldsymbol{w}_t, \phi_{j,t}) dF(\phi_{j,t})$ represents the probability of remaining in the industry for firm j at date t with the industry state $\boldsymbol{w}_t$.

Then we can rewrite (1) as:

$$\sum_{l \in \{0,1\}^{J_t}} \left[P(\boldsymbol{w}_{t+1} | \boldsymbol{w}_t, \boldsymbol{l}_t, \boldsymbol{i}_t(\boldsymbol{w}_t)) \prod_{j=1,\ldots,J_t} \varsigma_{j,t}(\boldsymbol{w}_t)^{l_{j,t}} (1-\varsigma_{j,t}(\boldsymbol{w}_t))^{1-l_{j,t}} \right] \tag{2}$$

Equation (2) implies that the probability of transition from state $\boldsymbol{w}_t$ to state $\boldsymbol{w}_{t+1}$ hinges on the probabilities of staying $\boldsymbol{\varsigma}_t(\boldsymbol{w}_t) = (\varsigma_{1,t}(\boldsymbol{w}_t), \ldots, \varsigma_{J_t,t}(\boldsymbol{w}_t))$. Thus, as mentioned by Dorazelski and Satterthwaite (2010), when forming an expectation over the future state of the industry, it is sufficient for a firm to know the probabilities of staying of its competitors $\boldsymbol{\varsigma}_{-j,t}(\boldsymbol{w}_t)$ instead of the decision rules $\boldsymbol{\chi}_{-j,t}(\boldsymbol{w}_t, \boldsymbol{\phi}_{-j,t})$.

Let us denote by $V_{j,t}(\boldsymbol{w}_t, \phi_{j,t})$ the expected net present value of all future profits for the firm j at date t, which is defined recursively by the solution of the Bellman equation:

$$\begin{aligned} V_{j,t}(\boldsymbol{w}_t, \phi_{j,t}) = \sup_{\substack{\chi_{j,t}(\boldsymbol{w}_t, \phi_{j,t}) \in \{0,1\} \\ i_{j,t}(\boldsymbol{w}_t)}} & \pi_{j,t}(\boldsymbol{w}_t) + (1-\chi_{j,t}(\boldsymbol{w}_t, \phi_{j,t}))\phi_{j,t} + \chi_{j,t}(\boldsymbol{w}_t, \phi_{j,t}) \\ & \times \left\{ -c(i_{j,t}(\boldsymbol{w}_t)) + \delta E\left\{ V_{j,t+1}(\boldsymbol{w}_{t+1}) | \boldsymbol{w}_t, i_{j,t}(\boldsymbol{w}_t), \boldsymbol{i}_{-j,t}(\boldsymbol{w}_t), \boldsymbol{\varsigma}_{-j,t}(\boldsymbol{w}_t) \right\} \right\} \end{aligned} \tag{3}$$

where δ represents the discount rate. The expression $V_{j,t}(\boldsymbol{w}_t, \phi_{j,t})$ denotes the value function after the firm has drawn its scrap value, and $V_{j,t}(\boldsymbol{w}_t) = \int V_{j,t}(\boldsymbol{w}_t, \phi_{j,t}) dF(\phi_{j,t})$ denotes the value function before the firm has drawn its scrap value. If the firm stays in the industry, it gets the profit from the product market competition plus the continuation value minus the cost of investment. If the firm leaves the industry, it gets the profit from the product market competition plus the scrap value of exit. By comparing the value of staying (i.e. the continuation value minus the cost of investment) with the

scrap value of exiting a firm takes the decision to exit i e to sell its assets to another firm in the industry) Note that it takes one time period to implement the exit decision For instance a firm that takes the exit decision at the beginning of period t gets the profit from product market competition and leaves effectively the market at the end of period t

From (3) the optimal decision of firm j at date t to remain in the industry is a cut off rule characterized by:

$$\chi_{j\,t}(w_t, \phi_{j\,t}) = \begin{cases} 1 & \text{if } \phi_{j\,t} < \bar{\phi}_{j\,t}(w_t) \\ 0 & \text{if } \phi_{j\,t} > \bar{\phi}_{j\,t}(w_t) \end{cases} \qquad (4)$$

where $\bar{\phi}_{j\,t}(w_t) = \sup_{i_{j\,t}(w_t)} \left(-c\left(i_{j\,t}(w_t)\right) + \delta E\left\{V_{j\,t+1}(w_{t+1} \mid w_t, i_{j\,t}(w_t), i_{-j\,t}(w_t), \varsigma_{-j\,t}(w_t))\right\}\right)$ Assuming that $\phi_{j\,t}$ follows a normal distribution $N(\phi, \omega)$ we can write (4) as:

$$\chi_{j\,t}(w_t, \phi_{j\,t}) = \begin{cases} 1 & \text{if } \dfrac{\phi_{j\,t} - \phi}{\omega} < \dfrac{\bar{\phi}_{j\,t}(w_t) - \phi}{\omega} \\ 0 & \text{otherwise} \end{cases} \qquad (5)$$

Thus the decision rule $\chi_{j\,t}(w_t, \phi_{j\,t}) = 1$ can be written equivalently $1(\phi_{j\,t} < \bar{\phi}_{j\,t}(w_t))$ or $1\left(\dfrac{\phi_{j\,t} - \phi}{\omega} < \dfrac{\bar{\phi}_{j\,t}(w_t) - \phi}{\omega}\right)$ where $1(\ )$ represents the indicator function This point is crucial when we derive the estimation algorithm to control for selection bias in the demand equation Then we can write:

$$\begin{aligned} \varsigma_{j\,t}(w_t) &= \int \chi_{j\,t}(w_t, \phi_{j\,t})\, dF(\phi_{j\,t}) \\ &= \int 1\left(\frac{\phi_{j\,t} - \phi}{\omega} < \frac{\bar{\phi}_{j\,t}(w_t) - \phi}{\omega}\right) dF(\phi_{j\,t}) \\ &= F\left(\frac{\bar{\phi}_{j\,t}(w_t) - \phi}{\omega}\right) \end{aligned} \qquad (6)$$

where $F(\ )$ is the cumulative distribution function of the standard normal distribution

Then we obtain the value of the threshold such that a firm stays or exits the market using the expression:

$$F^{-1}\left(\varsigma_{jt}(w_t)\right) = \frac{\bar{\phi}_{jt}(w_t) - \phi}{\omega} \qquad (7)$$

Assuming that we observe the state of the industry w_t in the data we are able to estimate the probabilities of remaining in the industry $\hat{\varsigma}_{jt}(w_t)$ for all (j, t) (see the next section for details on the estimation algorithm). Next inverting the standard normal distribution and evaluating at $\hat{\varsigma}_{jt}(w_t)$ we can compute the threshold value $(\bar{\phi}_{jt}(w_t) - \phi)/\omega$ This threshold will allow us to compute a correction term in order to take into account the selection bias in the estimation of the demand model The next section presents the demand model and the estimation algorithm to control for selection bias

4 Demand model and consumer surplus: econometric methodology

We first state the problem of selection bias by presenting a demand model that will be used to compute the consumer welfare for each time period between 1980 and 2006 Next we present the estimation algorithm that allows us to control for this selection effect using the dynamic model of firm behavior presented above We also explicitly discuss the issues linked to the exogenous process of the unobserved state variable in the demand model when a panel data is characterized by a self selection effect In some sense this discussion can be related to Wooldridge (1995) and Semykina and Wooldridge (2005) since we include firm fixed effects in the analysis

4.1 The demand framework

Following Berry (1994) we group different railroad firms into two groups and exclusive sets $g = 0, 1$ where $g = 0$ denotes the outside option and $g = 1$ denotes the group containing the railroad firms The utility of a consumer i from choosing the railroad firm j is:

$$u_{ijt} = \delta_{jt} + \zeta_{gt} + (1 - \sigma_g)\varepsilon_{ijt}$$

where δ_{jt} is the mean utility of choosing railroad j at time t and ε_{ijt} is identically and independently distributed extreme value The variable ζ_{gt} is common to all firms in group g and follows a Cardell (1997) distribution $C(\sigma)$ with $\sigma \in (0,1)$ The parameter σ represents the within group correlation of all the alternatives in the group $g = 1$

Regarding the mean utility $\delta_{jt} = x_{jt}\boldsymbol{\beta} + \theta k_{jt} - \alpha p_{jt} + \xi_{jt}$ where $\boldsymbol{x}_{jt}$ is a vector of demand related variables $k_{jt} = \ln(K_{jt})$ represents the logarithm of the ways and structures capital stock of the railroad firm j p_{jt} is the price of using the railroad firm j to provide the freight service and ξ_{jt} represents the unobservable efficiency of railroad firm j at time t

Before entering detailing the estimation algorithm it is important to understand the link between the demand model and the dynamic exit model We define the observed quality of firm j at time t as $\nu_{jt} \equiv x_{jt}\beta + \theta k_{jt}$ where ν_{jt} denotes the observed state variable for the firm j at date t in the dynamic model The second state variable ξ_{jt} which is unobservable to the econometrician is directly incorporated into the demand model These definitions of the state variables $\boldsymbol{w}_{jt} = (\nu_{jt}\ \xi_{jt})$ imply that the spot profit function can be written as $\pi_{jt}(J_t, \boldsymbol{w}_t)$ where $\boldsymbol{w}_t = (\boldsymbol{w}_{1t} \ldots \boldsymbol{w}_{jt} \ldots \boldsymbol{w}_{J_t t})$ This connection between the dynamic model and the demand model is essential to have a coherent framework

Using Berry (1994) and Train (1999) the formulas that characterize this nested logit model are:

$$s_{jt|g}(\delta\ \sigma) = \frac{\exp\left(\frac{\delta_{jt}}{1-\sigma}\right)}{\sum_{j\in g}\exp\left(\frac{\delta_{jt}}{1-\sigma}\right)}$$ for railroad firm j in group g at time t and

$$s_{gt}(\delta\ \sigma) = \frac{D_{gt}^{1-\sigma}}{\sum_{g} D_{gt}^{1-\sigma}}$$

where $s_{j\,t|g}(\cdot)$ denotes the within market share of firm j at time t in the group $g=1$, δ denotes the vector of mean utilities of all railroad firms, σ denotes the within group correlation of railroad firms and $D_{g\,t} \equiv \sum_{j\in g} \exp\left(\delta_{j\,t}/(1-\sigma)\right)$. Then, the market share of the outside alternative is given by $s_{0\,t}(\boldsymbol{\delta},\sigma)=1\Big/\sum_g D_{g\,t}^{1-\sigma}$ and the market share for the railroad firm j can be expressed as:

$$s_{j\,t}(\delta,\sigma)=s_{j\,t|g}(\delta,\sigma)s_{g\,t}(\delta,\sigma)=\frac{\exp\left(\dfrac{\delta_{j\,t}}{1-\sigma}\right)}{D_{g\,t}^{\sigma}\left[\sum_g D_{g\,t}^{1-\sigma}\right]} \tag{8}$$

The consumer welfare can thus be computed as:

$$CS_t=\frac{1}{\alpha}\ln\left(\sum_{g=0}^{1} D_{g\,t}{}^{1-\sigma}\right)=\frac{1}{\alpha}\ln\left(1+D_{1\,t}{}^{1-\sigma}\right)$$

Following Berry (1994), for a particular railroad firm j at year t, the estimating equations are:

$$\ln s_{j\,t}-\ln s_{0\,t}=x_{j\,t}\boldsymbol{\beta}+\theta k_{j\,t}-\alpha p_{j\,t}+\sigma\ln s_{j\,t|g}+\xi_{j\,t} \tag{9}$$

When we do not take into account the selection bias, we specify the following conditional moment restriction:

$$E\left[\xi_{j\,t}\middle|z_{j\,t}\right]=0 \tag{10}$$

for a set of instruments $z_{j\,t}$, and consistent estimates of the demand parameters are obtained through GMM using an optimal weight matrix. The instrumentation approach is standard in the literature. Instruments include the exogenous demand related variables that are included in the regression, the corresponding BLP (Berry, Levinsohn, and Pakes (1995)) instruments, and some cost shifter variables.[7]

However, as mentioned in the introduction, it is likely that there is selection on ξ due to firms with low values of ξ exiting. In this case, the moment condition (10) would in general be violated

[7] The variables used in the estimation are described in section 5.

This problem might be severe in our data since an important number of railroad firms exited the industry by selling their assets to other firms in the industry Therefore the condition (10) must be written as:[8]

$$E\left[\xi_{j\,t}\middle|z_{j\,t}\;\; r_{j\,t}=1\right]\neq 0 \tag{11}$$

where $r_{j\,t}$ is the selection indicator:

$$r_{j\,t}=\begin{cases}1 & \text{if firm } j \text{ is active at year } t\\ 0 & \text{otherwise}\end{cases} \tag{12}$$

Formally we write equation (9) as:

$$\ln s_{j\,t}-\ln s_{0\,t}=x_{j\,t}\beta+\theta k_{j\,t}-\alpha p_{j\,t}+\sigma\ln s_{j\,t|g}+E\left[\xi_{j\,t}\middle|z_{j\,t}\;\; r_{j\,t}=1\right]+e_{j\,t} \tag{13}$$

where $E\left[e_{j\,t}\middle|z_{j\,t}\;\; r_{j\,t}=1\right]=0$ by construction since $e_{j\,t}\equiv\xi_{j\,t}-E\left[\xi_{j\,t}\middle|z_{j\,t}\;\; r_{j\,t}=1\right]$

We now describes the estimation algorithm to correct for the selection bias in the demand equation (9) From equation (13) the correction term for the selection bias is $E\left[\xi_{j\,t}\middle|z_{j\,t}\;\; r_{j\,t}=1\right]$ From the dynamic model we know that it takes one time period to implement the exit decision Thus the firm is active at time t if it decides to stay on the market in the previous period $t-1$ that is if $\phi_{j\,t-1}<\bar{\phi}_{j\,t-1}(w_{t-1})$ Thus to be able to compute the correction term we need to condition with respect to the previous state of the industry w_{t-1} This leads to the following correction term:

$$\begin{aligned}E[\xi_{j\,t}\,|\,z_{j\,t}\;\; w_{t-1}\;\; r_{j\,t}=1]&=E[\xi_{j\,t}\,|\,z_{j\,t}\;\; w_{t-1}\;\; \chi_{j\,t-1}=1]\\&=E[\xi_{j\,t}\,|\,z_{j\,t}\;\; w_{t-1}\;\; \phi_{j\,t-1}<\bar{\phi}_{j\,t-1}(w_{t-1})]\end{aligned} \tag{14}$$

It means that in the presence of attrition the object of interest is $E[y_{j\,t}\,|\,z_{j\,t}w_{t-1}\;\; r_{j\,t}=1]$ where $y_{j\,t}$ denotes the dependent variable $\ln s_{j\,t}-\ln s_{0\,t}$

[8] Using the moment conditions in (10) might lead to inconsistent estimates of the demand parameters which will yield inconsistent estimates of the consumer welfare and inconsistent estimates of the marginal costs for firm j at time t (see Berry (1994) for the method to recover the marginal costs from the demand model)

Thus, the estimating equation can be written as follows:

$$y_{j,t} = \boldsymbol{x}_{j,t}\beta + \theta k_{j,t} - \alpha p_{j,t} + \sigma \ln s_{j,t|g} + E[\xi_{j,t} | z_{j,t}, \boldsymbol{w}_{t-1}, r_{j,t} = 1] + e_{j,t} \qquad (15)$$

where $e_{j,t} = \xi_{j,t} - E[\xi_{j,t} | z_{j,t}, \boldsymbol{w}_{t-1}, r_{j,t} = 1]$. The moment conditions are derived using $E[e_{j,t} | z_{j,t}, \boldsymbol{w}_{t-1}, r_{j,t} = 1] = 0$.

For the following, the correction term for attrition is computed using the law of iterated expectation:

$$E[\xi_{j,t} | z_{j,t}, \boldsymbol{w}_{t-1}, r_{j,t} = 1] = E[E[\xi_{j,t} | z_{j,t}, \boldsymbol{w}_{t-1}, \phi_{j,t-1}] | z_{j,t}, \boldsymbol{w}_{t-1}, r_{j,t} = 1] \qquad (16)$$

The exact form of the correction term depends on the specification of the process for the unobserved firm efficiency $\xi_{j,t}$. The following part of this section presents the estimation algorithm depending on the process of the unobserved firm efficiency $\xi_{j,t}$.

4.2 Estimation algorithm: firm fixed-effect and attrition

We assume the following specification for the unobserved efficiency:

$$\xi_{j,t} = c_j + \rho \phi_{j,t-1} + \tau_{j,t} \qquad (17)$$

where c_j denotes the firm fixed effect, $\phi_{j,t-1}$ denotes the scrap value, and $\tau_{j,t}$ denotes an error term independent from c_j and $\phi_{j,t-1}$.

The inclusion of a firm fixed effect can be related to the work of Wooldridge (1995) and Semykina and Wooldridge (2005). Indeed, incorporating firm fixed effect is attractive when we suspect firms to select out of the sample based on unobserved fixed heterogeneity. This can also be related to Nevo (2000, 2001) which advocates the use of fixed effect to ensure that the observed characteristics capture the true factors that determine utility and it improves the fit of the model.

The correction term can be written as:[9]

$$
\begin{aligned}
E[\xi_{j\,t} \mid z_j, w_{t-1}, r_j] &= E[E(\xi_{j\,t} \mid z_j, w_{t-1}, \phi_{j\,t-1}) \mid z_j, w_{t-1}, r_j] \\
&= c_j + \rho E[\phi_{j\,t-1} \mid w_{t-1}, r_j] \\
&= c_j + \rho E[\phi_{j\,t-1} \mid w_{t-1}, r_{j\,t} = 1] \\
&= c_j + \rho E[\phi_{j\,t-1} \mid w_{t-1}, \phi_{j\,t-1} < \overline{\phi}_{j\,t-1}(w_{t-1})] \\
&= c_j + \rho \lambda_j(w_{t-1}),
\end{aligned}
\qquad (18)
$$

where $\lambda_j(w_{t-1})$ denotes the mills ratio which is correcting for attrition, that is:

$$
\lambda_j(w_{t-1}) \equiv -\omega \frac{f\left(\dfrac{\overline{\phi}_{j\,t-1}(w_{t-1})}{\omega}\right)}{F\left(\dfrac{\overline{\phi}_{j\,t-1}(w_{t-1})}{\omega}\right)} \qquad (19)
$$

where $f(.)$ and $F(.)$ represents respectively the probability distribution function and the cumulative distribution function of the standard normal distribution, indeed, in the dynamic model, we have assumed that the scrap value $\phi_{j\,t-1}$ follows a normal distribution $N(0, \omega)$. Moreover, in equation (18) we have assumed that $E[\phi_{j\,t-1} \mid w_{t-1}, r_j] = E[\phi_{j\,t-1} \mid w_{t-1}, r_{j\,t} = 1]$. This is coherent with the dynamic exit model since the scrap value is independent over time.[10] Using (18) we see that the dynamic exit model provides a guidance for the construction of the correction term for attrition in a structural econometrics framework. It is like a Heckman two step procedure where the correction term comes from the economic theory using a dynamic model of exit behavior.

Then we obtain the equation:

$$
\ln s_{j\,t} - \ln s_{o\,t} = x_{j\,t}\beta + \theta k_{j\,t} - \alpha p_{j\,t} + \sigma \ln s_{j\,t|g=1} + c_j + \rho \lambda_{j\,t-1}(w_{t-1}) + e_{j\,t} \qquad (20)
$$

[9] Observe that we condition with respect to (z_j, r_j) and no more with respect to $(z_{j\,t}, r_{j\,t} = 1)$. This is due to the inclusion of the fixed effects in the analysis (see Wooldridge (1995, 2002, 2005)).

[10] Without this assumption, the term $E[\phi_{j\,t-1} \mid w_{t-1}, r_j]$ would be very complicated to compute (see Wooldridge (1995, 2002) and Semykina and Wooldridge (2005)).

where $E[e_{j\,t} \mid z_j \; \boldsymbol{w}_{t-1} \; \boldsymbol{r}_j] = 0$ by construction Note that we have normalized the variance of the scrap value $\omega = 1$ Indeed in equation (20) we can identify only the parameters $\rho\omega$ and not each parameter separately For the following we denote $\lambda_{j\,t-1}(\boldsymbol{w}_{t-1})$ by $\lambda_{j\,t-1}$ and $\ln s_{j\,t} - \ln s_{0\,t}$ by $y_{j\,t}$ We eliminate the firm fixed effect by a first difference analysis which leads to the following estimating equation:

$$\Delta y_{j\,t} = \Delta x_{j\,t}\beta + \theta\Delta k_{j\,t} - \alpha\Delta p_{j\,t} + \sigma\Delta\ln s_{j\,t|g=1} + \rho\Delta\lambda_{j\,t-1} + \Delta e_{j\,t} \qquad (21)$$

where Δ denotes the first difference operator and $\Delta e_{j\,t} = e_{j\,t} - e_{j\,t-1}$ for example It is worth to emphasize the choice of instruments for the equation (21) The obvious instruments are the exogenous variables $z_{j\,t}$ and $z_{j\,t-1}$ the exogenous demand variables $\boldsymbol{x}$ are included in the instruments $\boldsymbol{z}$) More details will be given in section 5 regarding the construction of the instruments $\boldsymbol{z}$

Let us now discuss the endogeneity of the other variables included in the estimating equation (21) The price denoted $p_{j\,t}$ and the within market share denoted $\ln s_{j\,t|g}$ are endogenous Thus the variables $\Delta p_{j\,t}$ and $\Delta\ln s_{j\,t|g=1}$ are also endogenous The discussion becomes a little more subtle for the variables $\Delta k_{j\,t}$ and $\Delta\lambda_{j\,t-1}$ Using the structure of the model we know that the variables $k_{j\,t}$ and $\lambda_{j\,t-1}$ are weakly exogenous Indeed $k_{j\,t} = \ln(K_{j\,t})$ and the capital stock is constructed using the relation $K_{j\,t} = K_{j\,t-1}(1-\delta) + I_{j\,t-1}$ where $I_{j\,t-1}$ represents the investment in the network at date $t-1$ (more details regarding the construction of the capital stock are available in section 5) From the dynamic model we know that the investment $I_{j\,t-1}$ is endogenous and it is a function of the previous state of the industry $\boldsymbol{w}_{t-1}$ This implies that the capital stock $K_{j\,t}$ and thus the proxy for network quality $k_{j\,t}$ are a function of $\boldsymbol{w}_{t-1}$ By construction the error term $e_{j\,t}$ in equation (20) is uncorrelated with the previous state of the industry $\boldsymbol{w}_{t-1}$ Thus the proxy for the network quality $k_{j\,t}$ is weakly exogenous since it is uncorrelated with the contemporaneous and the future error terms $e_{j\,s}$ $s \geq t$ and correlated with the past error term $e_{j\,s}$ $s \leq t-1$ This implies that in the estimating equation (21) the variable $\Delta k_{j\,t} = k_{j\,t} - k_{j\,t-1}$ is endogenous since $k_{j\,t}$ is correlated with $\Delta e_{j\,t}$ through $e_{j\,t-1}$

Nevertheless we can instrument $\Delta k_{j\,t}$ by using the $K_{j\,t-1}$ as instrument since the lag of the capital stock is a function of the state of the industry at date $t-2$ $\boldsymbol{w}_{t-2}$ and the error term is $\Delta e_{j\,t}$ is uncorrelated with the state of the industry at date $t-2$ for the estimation we have also added $K_{j\,t-2}$ as an instrument) Lastly we discuss the endogeneity of the first difference of the mills ratio $\Delta\lambda_{j\,t-1} = \lambda_{j\,t-1}\ \boldsymbol{w}_{t-1}) - \lambda_{j\,t-2}\ \boldsymbol{w}_{t-2})$ Like the stock of capital the mills ratio $\lambda_{j\,t-1}\ \boldsymbol{w}_{t-1})$ is also weakly exogenous since it is uncorrelated with $e_{j\,s}\ s \geq t$ and it is correlated with $e_{j\,s}\ s \leq t-1$ In the estimating equation 21) $\Delta\lambda_{j\,t-1}$ is endogenous since $\lambda_{j\,t-1}$ is correlated with $e_{j\,t-1}$ and thus with $\Delta e_{j\,t}$ We instrument $\Delta\lambda_{t-1} = \lambda_{j\,t-1} - \lambda_{j\,t-2}$ by the second lag of the mills ratio that is $\lambda_{j\,t-2}$

To summarize the choice of the instruments is guided by the structure of the model Hence during the estimation accepting the over identifying restriction may be interpreted as accepting the structure of the model as well We will come back on this issue when we present the estimation results in section 6

In the estimating equation 21) we have assumed that we know the previous state of the industry $\boldsymbol{w}_{t-2}$ since we use the conditioning $E\ \Delta e_{j\,t} \mid \boldsymbol{z}_j\ \boldsymbol{w}_{t-2}\ \boldsymbol{r}_j] = 0$ To make the estimation feasible we need to use the following iterative algorithm:

1) Start with an initial guess of the vector of demand parameters denoted $\widehat{\boldsymbol{\mu}} = \ \widehat{\boldsymbol{\beta}}\ \hat{\theta}\ \hat{\alpha}\ \hat{\sigma})$
2) Using the equation 9) we compute an estimate of the unobserved state variable that represents the unobserved firm efficiency $\hat{\xi}_{j\,t}$ and we compute the observed state variable of each firm $\hat{\nu}_{j\,t} = x_{j\,t}\widehat{\boldsymbol{\beta}} + \hat{\theta}k_{j\,t}\ \ \forall\ j\ t)$
3) We compute the probabilities of remaining in the industry as a function of the industry state $\hat{\varsigma}_{j\,t}\ \widehat{\boldsymbol{w}}_t)$ where $\widehat{\boldsymbol{w}}_t = \ J_t\ \hat{\nu}_{j\,t}\ \hat{\boldsymbol{\nu}}_{-j\,t}\ \hat{\xi}_{j\,t}\ \hat{\boldsymbol{\xi}}_{-j\,t})$ using a probit model and $\hat{\boldsymbol{\nu}}_{-j\,t}$ and $\hat{\boldsymbol{\xi}}_{-j\,t}$ represent respectively the sum of the observed and the unobserved state variable for the competitors

4) The threshold value $\frac{\bar{\phi}_{j\,t-1}(\hat{w}_{t-1})}{\omega} = F^{-1}(\hat{\varsigma}_{j\,t-1}(\hat{w}_{t-1}))$ is computed and we obtain the mills ratio $\hat{\lambda}_{j\,t-1}(\hat{w}_{t-1})$ as a correction term for attrition (see equation (19)) We are also able to recover $\hat{\lambda}_{j\,t-2}(\hat{w}_{t-2})$

5) We estimate the regression (21) by an instrumental variable regression using the instruments $z_{j\,t}$ $z_{j\,t-1}$ $K_{j\,t-1}$ $K_{j\,t-2}$ and $\hat{\lambda}_{j\,t-2}$

6) Using the new demand estimates $\hat{\mu} = (\hat{\beta}\ \hat{\theta}\ \hat{\alpha}\ \hat{\sigma})$ we repeat steps 2 5 until convergence of the demand estimates

The key point in the estimation algorithm is the assumption on the exogenous process of the unobserved firm efficiency $\xi_{j\,t}$ which is necessary to compute the correction term in equations (20) and (21) We are particularly concerned with the presence of serial correlation This would have an impact on the correction term $E[\xi_{j\,t} | z_j, w_{t-1}\ r_j] = E[E(\xi_{j\,t} | z_j, w_{t-1}\ \phi_{j\,t-1}) | z_j\ w_{t-1}\ r_j]$ since $\xi_{j\,t-1}$ is part of the previous state of the industry w_{t-1} To deal with this issue the next subsection includes explicitly serial correlation in the unobserved firm efficiency

4.3 Estimation algorithm: firm fixed-effect, serial correlation, and attrition

We assume the following exogenous process for the unobserved efficiency:

$$\xi_{j\,t} = c_j + \gamma\xi_{j\,t-1} + \rho\phi_{j\,t-1} + \tau_{j\,t} \qquad (22)$$

where the error term $\tau_{j\,t}$ is identically and independently distributed The correction term can be written as:

$$\begin{aligned} E[\xi_{j\,t} | z_j, w_{t-1}\ r_j] &= E[E(\xi_{j\,t} | z_j, w_{t-1}\ \phi_{j\,t-1}) | z_j\ w_{t-1}\ r_j] \\ &= c_j + \gamma\xi_{j\,t-1} + \rho\lambda_j(w_{t-1}) \end{aligned} \qquad (23)$$

The equation (20) becomes:

$$\ln s_{j\,t} - \ln s_{o\,t} = x_{j\,t}\beta + \theta k_{j\,t} - \alpha p_{j\,t} + \sigma \ln s_{j\,t|g=1} + c_j + \gamma\xi_{j\,t-1} + \rho\lambda_{j\,t-1} + e_{j\,t} \qquad (24)$$

We eliminate the fixed effect by a first difference which leads to the following estimating equation:

$$\Delta y_{j\,t} = \Delta x_{j\,t}\beta + \theta \Delta k_{j\,t} - \alpha \Delta p_{j\,t} + \sigma \Delta \ln s_{j\,t|g=1} + \gamma \Delta \xi_{j\,t-1} + \rho \Delta \lambda_{j\,t-1} + \Delta e_{j\,t} \quad (25)$$

where $E[\Delta e_{j\,t} | z_j, r_j, w_{t-2}] = 0$ by construction. The instruments used are the same as in the previous subsection. We just need to add an instrument regarding the variable $\Delta \xi_{j\,t-1} = \xi_{j\,t-1} - \xi_{j\,t-2}$. This variable is endogenous since $\xi_{j\,t-1}$ is correlated with $e_{j\,t-1}$ and thus with $\Delta e_{j\,t}$. This variable is instrumented by $\hat{\xi}_{j\,t-2}$ like in the previous subsection, the second step of our iterative algorithm allows to recover previous values of the unobserved efficiency $\hat{\xi}_{j\,s}$, $s \leq t-1$, before we estimate the equation (25). Again, the choice of this instrument is guided by the structure of the model.

4.4 Estimation algorithm: no attrition

The purpose of this section is to show that if attrition does not matter, then our iterative algorithm vanishes to a simple instrumental variable estimation.

Assume the following process for the unobserved firm efficiency:

$$\xi_{j\,t} = c_j + \gamma \xi_{j\,t-1} + \tau_{j\,t} \quad (26)$$

Then, the correction term is equal to:

$$E[\xi_{j\,t} | z_j, w_{t-1}, r_j] = c_j + \gamma \xi_{j\,t-1} \quad (27)$$

which implies the following estimating equation:

$$y_{j\,t} = x_{j\,t}\beta + \theta k_{j\,t} - \alpha p_{j\,t} + \sigma \ln s_{j\,t|g=1} + c_j + \gamma \xi_{j\,t-1} + e_{j\,t} \quad (28)$$

where $E[e_{j\,t} | z_j, w_{t-1}, r_j] = 0$ by construction and $y_{j\,t} \equiv \ln s_{j\,t} - \ln s_{0\,t}$. In other words, the object of interest can be defined as $E[y_{j\,t} | z_j, w_{t-1}, r_j] = 0$. Since we do not include a mills ratio to correct for

attrition it is not necessary to condition with respect to the previous state of the industry $\boldsymbol{w}_{t-1}$ Thus we consider the following object:

$$E[y_{j\,t} \mid z_j\ \boldsymbol{r}_j] = E[E[y_{j\,t} \mid z_j\ \boldsymbol{r}_j\ \boldsymbol{w}_{t-1}] \mid z_j\ \boldsymbol{r}_j] \tag{29}$$

Replacing $y_{j\,t}$ by its expression in (28) we obtain:

$$\begin{aligned} E[y_{j\,t} \mid z_j\ \boldsymbol{r}_j] &= E[x_{j\,t}\beta + \theta k_{j\,t} - \alpha p_{j\,t} + \sigma \ln s_{j\,t|g=1} + c_j + \gamma \xi_{j\,t-1} \mid z_j\ \boldsymbol{r}_j] \\ &= x_{j\,t}\beta + \theta k_{j\,t} - \alpha p_{j\,t} + \sigma \ln s_{j\,t|g=1} + c_j + \gamma E[\xi_{j\,t-1} \mid z_j\ \boldsymbol{r}_j] \\ &= x_{j\,t}\beta + \theta k_{j\,t} - \alpha p_{j\,t} + \sigma \ln s_{j\,t|g=1} + c_j \end{aligned} \tag{30}$$

where we assume that $E[\xi_{j\,t-1} \mid z_j\ \boldsymbol{r}_j] = 0$ The estimation of equation (30) is standard and our iterative algorithm vanishes to a standard regression We estimate it by first differencing to remove the firm fixed effect and we use the same instruments as in the previous subsections More precisely we use the exogenous variables $z_{j\,t}$ and $z_{j\,t-1}$ We also use the previous capital stock $K_{j\,t-1}$ and $K_{j\,t-2}$ as instruments

The last case that we consider is:

$$\xi_{j\,t} = c_j + \tau_{j\,t} \tag{31}$$

This implies that we do not consider attrition and serial correlation in the exogenous process of the unobserved firm efficiency With respect to the previous case we add the current capital stock $K_{j\,t}$ as an instrument Indeed if we assume that there is not serial correlation and no attrition in the process of the unobserved firm efficiency the structure of the model tells us to include this variable as an instrument

5 Data

The main sources of data are the “Analysis of Class1 Railroads” hereafter “Analysis”) published annually by the Association of American Railroads (AAR) The “Analysis” is based on regulatory reports that railroads submit to the Surface Transportation Board (STB) In order to adjust

for the effect of inflation we convert the monetary variables in current dollars $1982) using the Consumer Price Index from the Statistical Abstract of the US see also the US Bureau of Labor Statistics) We focus on a panel of 42 Class1 rail companies that operated in US between 1980 and 2006 These 42 Class 1 firms are defined as accounting entities[11] Figure 1 and Figure 2 list all the takeovers that happened in the railroad industry We define a takeover between two firms such that one firm buys another firm There are two elements of ambiguity for the construction of the merged entities namely the merged firms *CSX* and *NS* in 1986 These two firms appear in 1986 and are the results of the mergers of several firms We assume that the merger parties have sold their assets to the firm with the highest market share before the merger[12] Thus we assume that the firms *BO* and *CO* have sold their assets to *SBD* in 1986 and the firm *NW* has sold its assets to *SOU* in 1986 This treatment of merger yields an unbalanced panel data with an attrition characteristic such that see Wooldridge 2002 Chapter 17):

$$r_{j\,t} = 1 \Rightarrow r_{j\,\tau} = 1 \text{ for all } \tau \leq t-1$$

Regarding the construction of the price of providing freight services we build the series in the ton miles unit In particular for each firm j active in year t the “Analysis” gives the *Total Gross Freight Revenue* line 599) and the *Total Ton-Miles* line 711) We compute the price of freight in ton miles using the formula:

$$p_{j\,t} = \frac{Total\ gross\ freight\ revenue\ of\ firm\ j\ at\ year\ t}{Total\ ton\ \ miles\ of\ firm\ j\ at\ year\ t}$$

This allows us to build price series that are consistent with the study of the Surface Transportation Board STB) “Study of Railroad Rates: 1985 2007” 2009) using the data from the “Analysis of Class1 Railroads” 1980 2006) Indeed the STB has access to confidential and very detailed data in particular the *Official Waybill Sample* that records the prices of the commodities

[11] Firms that lost their Class1 status during 1978 2006 are excluded from the data These firms did not really exit from the industry but they disappeared from the Class 1 railroads due to changes in the rule that defines a Class 1 railroad firm Currently the Surface Transportation Board STB) defines a Class I railroad in the United States as having annual carrier operating revenues of $346 8 in 2006

[12] This assumption reflects what we observe in the data for all the railroad firms in Figure 1 and Figure 2

shipped in the US) whereas we have access to the "Analysis of Class 1 Railroads" 1980 2006) where pricing information is not directly available For a particular year t the industry price index is computed by a weighted average of the prices $p_{j\,t}$ of active firms where the weights are equal to the market share of firm j at year t:

$$s_{j\,t|g=1} = \frac{Total\ ton\ \ miles\ of\ firm\, j\ at\ year\ t}{Total\ industry\ ton\ \ miles\ at\ year\ t}$$

We compute two industry price indices namely one price index where the *Total Gross Freight Revenue* is in current dollars and another price index where it is in $1982 using the Consumer Price Index as a deflator see the Statistical Abstract of the US) In

Figure 5 we see that the evolution of our price index is consistent with the evolution of the price index built by the Surface Transportation Board 2009 see Figure 6 in Appendix) Thus using ton miles allows us to consistently reproduce the evolution of railroad rates displayed in the railroad rates study of the Surface Transportation Board 2009) Secondly using data from the US Department of Transport Bureau of Transportation Statistics) we are able to obtain the total size of the freight market in the US that is the freight provided by air truck railroad water and pipeline) in the ton miles unit Thus we can construct the market share of each railroad firm and the market share of the outside alternative) by not making an arbitrary assumption about the total size of the freight market Figure 4 shows the evolution of the Class 1 railroad freight market share during the period 1980 2006 the market considered is the US national freight market) We see that the market share of the Class 1 US railroad firms has increased from 20% in 1980 to 30% in 2006

The construction of the capital stock follows the methodology of Berndt Braeutigam Friedlaender McCullough and Meyer 1992) Accordingly we start from an authoritative estimate of the reproduction cost of capital in 1973 using Nelson 1975) and update the stock of capital of firm j using the perpetual inventory relation:

$$K_{j\,t+1} = K_{j\,t}\ 1 - d) + I_{j\,t} \qquad 32)$$

where I_{jt} represents the real investment in $1982) at year t The depreciation rate d is derived by solving an equation that allows railroad capital to depreciate exponentially over 25 years to a salvage value of 10 percent[13] The "Analysis of Class1 Railroads 1980 2006)" allows the measurement of the nominal investment which is then converted into real value $1982) The main difficulty consists in measuring this nominal investment component for way and structures capital Before 1982 railroads used "betterment" accounting in which the work on railroad way and structures is listed as an expense and thus excluded from the undepreciated book value of road line 67 in the "Analysis") Thus a first difference of the undepreciated book value of road allows measuring the nominal investment at every year After 1982 the railroad industry adopted a depreciation accounting system where the work on way and structures is added to the book value of road It is thus necessary to remove the expenditures linked to the maintenance of the network line 174 minus line 172 in the "Analysis") from the undepreciated book value of road and then do a first difference to obtain the nominal investment This perpetual inventory process is iterated to bring the series of way and structure capital until 2006[14]

We follow Vellturo 1989) and Berndt et al 1993a 1993b) in using a set of exogenous demand related variables that can be constructed on a firm specific basis These variables —*coal consumption CCON) coal production CPRO) new car registrations NEWCAR) state population SPOP) oil prices OILP) farm income FARM) and value of shipment from manufacturing SHIPMENT)*— are measured on a state by state basis and then aggregated across states to be railroad specific and to conform to each railroads' operating territory These aggregations vary from year to year as some railroad firms exit the industry and some other railroad firms extend their networks by

[13] The 25 year assumption is based on Berndt et al 1992)

[14] It is worth to mention the treatment of takeovers in the construction of the capital stock For example consider the takeover between "UP" and "MKT" in 1987 see Figure 1) The "Analysis" gives us the data on the capital stock at the end of 1987 for "MKT" and "UP" and the data for the capital stock at the end of 1988 for the merged firm "UP MKT" To measure the investment of the merged firm "UP MKT" in 1988 it is necessary to know its capital stock at the beginning of 1988 However this data is not available in the "Analysis" This data exists in the initial R1 reports filled by the railroad firms in 1988 but the R1 reports for the period 1978 1995 are no more available except on microfiche in the library of the Surface Transportation Board in Washington DC Only the R1 reports for the period 1996 2006 are available on the website of the STB) Thus we make the arbitrary assumption that the capital stock of the merged firm "UP MKT" at the beginning of 1988 is eq11 4422 T) 2006 o) 18

buying the assets of the firms that exit the industry These variables are based on annual data from the Association of American Railroads the Department of Transport Statistics the US Energy Information Administration the US Department of Agriculture the US Department of Commerce with the “Annual Survey of Manufacturers” and the different economic censuses the US Federal Highway Administration and the Statistical Abstract of the US see Table 3 in Appendix for more details) In the next section the estimation results are reported by using only the *coal consumption (CCON)* variable this variable appeared to be the most important from a demand side perspective see also section 2 for a justification) Lastly we include a *quadratic time trend* in order to capture any unexplained productivity growth The strictly exogenous variables include *the coal consumption CCON*) the *quadratic time trend* the BLP instrument of the coal consumption [15] the *miles of road operated ROAD*) *the average length of haul HAUL*) the lag of these six variables and the lag of the BLP instrument for *ROAD* The two variables *ROAD* and *HAUL* are used as instruments since they are considered as cost shifters see Berndt et al 1993a 1993b) Ivaldi and McCullough 2001 2008)) The use of the *capital stock* as instrument is already mentioned in section 4 Table 4 in Appendix) reports the descriptive statistics

6 Estimation results

This section presents the estimation results of the demand model using our algorithm to correct for selection due to attrition of firms The crucial point in the estimation method to correct for selection bias is to derive a cut off value denoted $\bar{\phi}_{j\,t}$ that determines the exit staying decision [16] The equations 6) and 7) are the bases of the estimation strategy Indeed in the estimation algorithm we have to compute only the probability of staying in the industry in order to correct for the selection bias

[15] The construction of the BLP instruments comes from Berry Levinsohn and Pakes 1995) For example for a railroad firm j the BLP instrument of the exogenous variable x_j is $\sum_{r \neq j} x_r$ where r denotes a competitor of the railroad firm j

[16] Olley and Pakes 1996) use a similar idea in the context of production function with a constant scrap value over time and no strategic interactions between firms Their threshold is determined as a cut off value in term of productivity

due to exit The advantage of our methodology is to correct for the attrition bias using a dynamic model to derive a selection equation but we do not need to solve the dynamic model to obtain the correction term The method is relatively simple and the only computational burden might concern the iterative process but convergence was very quick to achieve during the estimations

Table 5 report the estimation results using four different methodologies:

- In column 1 we implement a standard fixed effect analysis with no serial correlation and no attrition see equation 31))
- In column 2 we implement a fixed effect analysis by correcting for the selection bias see equation 17))
- In column 3 we implement a fixed effect analysis by taking into account serial correlation but not attrition see equation 26))
- In column 4 we implement a fixed effect analysis by correcting for attrition and by taking into account the potential serial correlation in the unobserved firm efficiency see equation 22))

The results from column 1 in Table 5 is puzzling since it implies that an increase in the quality of a network has a negative impact on the utility of the consumers the coefficient of the capital stock is negative $-0\ 0645$ which is counter intuitive) This negative estimate may come from a selection bias as Olley and Pakes 1996) mentioned in their paper on production function As column 2 shows correcting for selection implies a positive impact of the capital stock on demand $+0\ 0910$) The intuition of this result follows from 18) We know that $E\ \phi_{j\,t-1} | \phi_{j\,t-1} < \bar{\phi}_{j\,t-1}\ w_{t-1})$) is equal to $-f\ \bar{\phi}_{j\,t-1}) \big/ F\ \bar{\phi}_{j\,t-1})$ where we normalized the variance ω of the scrap value to one We can check that this expectation is negative for $\bar{\phi}_j$ positive and this is the case in our model Indeed we recover the threshold $\bar{\phi}_j$ by evaluating the inverse of the normal cumulative distribution function in the probability of staying in the industry since the estimated probabilities denoted $\hat{\zeta}$ are always above $0\ 5$ it implies that the threshold are positive $\bar{\phi}_{j\,t} > 0\ \forall\ j\ t)$ Moreover we know that $E\ \xi_{j\,t} | z_j\ w_{t-1}\ r_j) = c_j + \rho E\ \phi_{j\,t-1} | z_j\ w_{t-1}\ \phi_{j\,t-1} < \bar{\phi}_{j\,t-1}) = c_j + \rho \lambda_{j\,t-1}$ where the parameter ρ represents

the correlation between the unobserved state variable $\xi_{j\,t}$ and the value of exit $\phi_{j\,t}$ We expect $\rho > 0$ since it is reasonable to think that a higher value of exit is associated with a higher value of the unobserved state variable This implies that the correction term for attrition is negatively correlated with the stock of capital and we have a negative bias when we do not take into account the selection effect in column 1 [17] When we control for attrition in column 2 our estimation algorithm allows removing this negative bias and we find a positive impact of the quality of the network on the utility of the consumers +0 0910 We can also illustrate the selection bias by comparing column 3 and column 4 In the presence of serial correlation when we do not take into account the attrition we underestimate by 50% the coefficient of the network quality 0 07 in column 3 instead of 0 14 in column 4

Comparing column 1 and column 2 we also see that the price coefficient is under estimated when we do not control for attrition 38 91 in column 1 instead of 44 44 in column 2 When we control only for serial correlation column 3 we see that the price coefficient is almost the same 39 5622 as in column 1 We can interpret this result by saying that attrition is the main cause of the bias in the price coefficient α

To summarize by comparing columns 1 2 and 3 we see that attrition may create some important biases in the coefficient estimates which might lead to counter intuitive results Our estimation algorithm controls for the attrition selection bias and column 2 shows that attrition is significant at a 5% level However as mentioned in the subsection 4 3 some bias can remain in the estimates of column 2 due to serial correlation in the unobserved firm efficiency For example we can

[17] Intuitively increasing the proxy for network quality leads to an increase in the probability of staying This implies an increase in the threshold value of staying $\bar{\phi}_j$ and thus a decrease in the mills ratio λ_j Since ρ is positive it implies that the correction term for attrition $\rho\lambda_j$ is also decreasing Thus the correction term for attrition is negatively correlated with the proxy for network quality and this causes a negative bias in column 1 regarding the coefficient of the network quality The same reasoning applies for the coefficients of the variables x_j since they are included in the observed state variable $v_{j\,t} = x_{j\,t}\beta + \theta k_{j\,t}$ However in Table 5 we see that the coefficients of the coal consumption and of the quadratic time trend are roughly constant across the four specifications

prove that the coefficient of the correction term ρ is under estimated when we do not take into account the potential serial correlation the reasoning is similar to the previous one regarding the negative bias for the coefficient of the network quality This is coherent with column 4 in Table 5 which shows an increase in the coefficient of the correction term for attrition 0 7450 The price coefficient 56 7705 and the coefficient of the network quality 0 1446 are also increasing We also remark that the coefficient of the within market share denoted σ is decreasing from column 1 0 8582 to column 4 0 6639

Lastly the Sargan test is the best for the column 4 when we control for attrition and serial correlation This means that we accept the over identifying restrictions for column 4 and in some sense it also means that we accept the structure of the model see section 4 By comparing only the columns 1 2 and 3 we also see that the best Sargan test corresponds to column 2 where we control for attrition

We also report the estimations of the exit staying policy function in Table 6 This was useful in the algorithm to compute the probability of staying in the industry and thus the correction term for selection bias As it is mentioned in the dynamic model the staying exit policy function depends on the industry state $w_t = J_t, w_{1t} \dots w_{jt} \dots w_{J_t t}$ where $w_{jt} = \nu_{jt} \ \xi_{jt}$ Thus for a particular time period t we define the policy function of firm j as a function of the number of active firms J_t the own observed state variable ν_{jt} the observed state variables of the competitors $\sum_{l \neq j}^{J_t} \nu_{lt}$ the own unobserved state variable ξ_{jt} and the unobserved state variables of the competitors $\sum_{l \neq j}^{J_t} \xi_{lt}$ We also add a quadratic time trend The parameter estimates have the correct signs An increase in the own observed and unobserved state variables is likely to increase the probability of staying in the industry and an increase in the state variables of the competitors is likely to decrease the probability of staying Lastly when more firms are active the possibilities for takeovers are also increasing thus a

particular railroad firm is more likely to exit the market by selling its assets to another firms This is likely to decrease the probability of staying in the market

Next we report the evolutions of the consumer surplus and the evolution of the marginal costs of firms using the parameter estimates of column 4 We choose this particular model since it controls for attrition and serial correlation in the unobserved firm efficiency Moreover the Sargan test is the best and does not reject the over identifying restrictions see Table 5)

As it is standard in the literature we recover the marginal costs for each firm at a particular year by assuming that the firms compete in Bertrand Using the first order condition

$$(p_{jt} - mc_{jt}) \frac{\partial s_{jt}}{\partial p_{jt}} + s_{jt} = 0$$

we obtain the following formula for the mark up:

$$p_{jt} - mc_{jt} = \frac{1-\sigma}{\alpha(1-\sigma s_{jt|g=1} - s_{jt}(1-\sigma))} \qquad (33)$$

Then we can recover the marginal costs from the estimates of the mark ups using the formula $mc_{jt} = -\widehat{(p_{jt} - mc_{jt})} + p_{jt}$ Table 7 reports some descriptive statistics about the mark ups and marginal costs Note that the specification of the demand model implies positive marginal costs for each firm Then for each year we compute an average marginal cost where the marginal cost of each firm is weighted by its within market share We do a similar procedure to construct an index for the industry markup Figure 7 and Figure 8 show the evolutions of these indexes for the industry marginal cost and for the industry markup

We are also able to recover the consumer surplus for each time period using the formula:

$$CS_t = \frac{1}{\alpha} \ln\left(1 + D_t^{1-\sigma}\right) \qquad (34)$$

where $D_t = \sum_{j=1}^{J_t} \exp\left(\frac{\delta_{j\,t}}{1-\sigma}\right)$ and $\delta_{j\,t} = x_{j\,t}\beta + \theta k_{j\,t} - \alpha p_{j\,t} + \xi_{j\,t}$ Figure 9 reports the evolution of consumer surplus over time

Using the evolution of the freight prices Figure 5) the evolution of the industry marginal costs Figure 7) the evolution of the industry's mark ups Figure 8) and the evolution of the consumer surplus over time Figure 9) we are able to see how the railroad firms have passed the efficiency gains due to takeovers to the final consumers

From a cost perspective the industry concentration led to a decrease of the marginal costs 80%) for the period 1980 2004 This is consistent with the literature on cost function Berndt et al 1993a 1993b) Ivaldi and McCullough 2001 2008)) Moreover Dennis et al 2010) mention other efficiency gains of takeovers leading to a more efficient use of equipment by rerouting traffic to avoid congestion on routes reduction of ton miles by choosing the shortest routes and routes dedicated to specialized services which might decrease the marginal costs as well

Moreover the consumer surplus increased by 60% during the period 1980 2004 see Figure 9 and Table 8) This is mainly due to two elements First the railroad firms passed part of the decrease in marginal costs to the consumers by decreasing the final prices 60%) and this decrease in the final prices had a positive impact on the consumer surplus The second element is the increase in the efficiency index of the industry 0 25 billion $1982 from 1980 to 2006 that is 18 80% see Figure 10) due to the concentration of the industry [18] This means that the takeovers have led to a reallocation of assets from the less efficient firms to the most efficient firms which has increased the efficiency of the industry [19] Moreover this increase in the industry efficiency led to an increase in the consumer welfare during the period 1980 2004 Using the estimates of Table 9 we compute the elasticities of the

[18] We construct this index by computing a weighted average where each recovered firm efficiency $\xi_{j\,t}$ is weighted by the within market share

[19] Since the recovered efficiency index is negative it means that the inefficiency has decreased over time

consumer welfare with respect to the efficiency index and the price index of the industry We obtain that an increase of 1% of the industry efficiency implies on average an increase of 1 06% of the consumer welfare and a decrease of 1% of the freight price implies on average an increase of the consumer welfare by 0 90% for the period 1980 2006

To summarize the estimates yield that the takeover waves have led to a decrease in prices due to lower marginal costs and to an increase in the industry efficiency due to reallocation of assets through takeovers Moreover the impact of the industry efficiency on the consumer welfare is at least as significant as the impact of prices

Regarding the mark ups Figure 8 shows the evolution of mark ups over time [20] From 1980 to 1999 Figure 8 indicates that the takeovers were profitable: the industry mark up increased by 20% However starting from 1999 we observe that the industry mark up has stopped to rise Interestingly this corresponds to the end of the takeover waves For instance in 1999 the Surface Transportation Board prevented a merger between Burlington Northern and Canadian National

Lastly from 2004 to 2006 we observe an increase in the final prices by 11 30% This raised a debate in the US regarding the captivity of shippers and the abuse of market power by the railroad firms However we find that the industry marginal cost increased by 26 80% during the same period This means that the increase in final prices was mainly due to an increase in the marginal costs rather than an abuse of market power

7 Conclusion

In this paper we analyzed the impact of the takeovers in the US rail freight industry on the marginal costs prices and consumer welfare We find that the concentration led to a decrease in marginal costs by 80% which was translated into lower prices by 60% This is coherent with the

[20] The construction of the index for the industry mark up is similar to the price and marginal cost indexes

literature on the estimation of cost function in the US railroad industry which finds important economies of density There could also be other efficiency gains of takeovers leading to more efficient use of equipment by rerouting traffic to avoid congestion on routes reduction of ton miles by choosing the shortest routes and routes dedicated to specialized services which might decrease the marginal costs as well However for the period 2004 2006 we find that the industry marginal cost increased by 26 80% and this was passed to the consumers through a price increase of only 11 3% This means that the price increase of freight services after 2004 was mainly due to an increase of marginal costs and not to an abuse of market power by the railroad firms Moreover during the period 1980 2006 the concentration of the industry led to a reallocation of assets from the less efficient firms to the more efficient firms This increase in industry efficiency also positively impacted the consumer welfare Overall the decrease in prices and the increase in the industry efficiency led an increase of the consumer welfare by 60% during the period 1980 2004 As a next step it would be interesting to analyze the data up to today to see if the evolutions of the marginal costs freight prices mark ups and consumer surplus have the similar trend as in the period 2004 2006

From a technical point of view the concentration of the industry led to an unbalanced panel characterized by attrition due to the selection of the most efficient firms We propose an estimation algorithm to deal with this issue in the context of a structural demand model Our estimation methodology corrects for selection by using a selection equation that is derived from a dynamic model that allows for the equilibrium force of exit We find that our methodology correcting for selection bias gives more plausible parameter estimates than the standard methodologies We believe that it is an important step since the estimation of demand model is often one of the most important ingredients in the empirical industrial organization literature

REFERENCES

Aguirregabiria Victor and Pedro Mira "Dynamic Discrete Choice Structural Models: A Survey " *Journal of Econometrics* 2009: 1 30

Aguirregabiria Victor and Pedro Mira "Sequential Estimation of Dynamic Discrete Games " *Econometrica* 75 no 1 2007 : 1 53

Aguirregabiria Victor and Pedro Mira "Swapping the Nested Fixed Point Algorithm: A Class of Estimators for Discrete Markov Decision Models " *Econometrica* 70 no 4 2002 : 1519 1543

Cardell N Scott "Variance Components Structures for the Extreme Value and Logistic Distributions with Application to Models of Heterogeneity " *Econometric Theory* 13 1997 : 1985 213

Chamberlain Gary "Analysis of Covariance with Qualitative Data " *Review of Economic Studies* 47 1980 : 225 238

Chamberlain Gary "Multivariate Regression Models for Panel Data " *Journal of Econometrics* 18 1982 : 5 46

Christensen Associates "A Study of Competition in the U S Freight Railroad Industry " *Surface Transportation Board* November 2008

Cohen Andrew and Mazzeo Michael "Market Structure and Competition Among Retail Depository Institutions " *Review of Economics and Statistics* 89 no 1 2007 : 60 74

Dennis Scott Vikram Maheshri and Clifford Winston "Long run effects of mergers: the case of U S Western Railroads " *Brookings (Working Paper)* March 2010

Dorazelski Ulrich and Mark Satterthwaite "Computable Markov Perfect Industry Dynamics " *RAND Journal of Economics* 41 no 2 2010 : 215 243

Dubois Pierre and Celine Bonnet "Inference on Vertical Contracts between Manufacturers and Retailers Allowing for Non Linear Pricing and Resale Price Maintenance " *RAND Journal of Economics* 41 no 1 2010 : 139 164

Ericson Richard and Ariel Pakes "Markov Perfect Industry Dynamics: A Framework for Empirical Work " *Review of Economics Studies* 62 1995 : 53 82

Friedlaender Ann F "Coal Rates and Revenue Adequacy in a Quasi Regulated Rail Industry " *RAND Journal of Economics* 23 no 3 1992 : 376 394

Government Accountability Office "Freight Railroads: Industry Health Has Improved but Concerns about Competition and Capacity should be Addressed " October 2006

Greene William H *Econometric Analysis.* Fifth Edition Prentice Hall 2002

Ivaldi Marc and Frank Verboven "Quantifying the Effects from Horizontal Mergers in European Competiton Policy " *International Journal of Industrial Organization* 23 2005 : 669 691

Ivaldi Marc and Gerard J McCullough "Density and Integration Effects on Class 1 U S Freight Railroads " *Journal of Regulatory Economics* 19 2001 : 161 182

Ivaldi Marc and Gerard J McCullough "Subadditivity Tests for Network Separation with an Application to U S Railroads " *Review of Network Economics* 7 2008 : 159 171

Ivaldi Marc and Gerard J McCullough "Welfare Tradeoffs in U S Rail Mergers " *Working Paper Toulouse School of Economics and University of Minnesota* 2005

Jovanovic Boyan and Peter L Rousseau "The Q Theory of Mergers " *American Economic Review Papers and Proceedings* 92 no 2 2002 : 198 204

Maksimovic Vojislav and Gordon Phillips "The Market for Corporate Assets: Who Engages in Mergers and Asset Sales and Are There Efficiency Gains?" *Journal of Finance* 56 no 6 2001 : 2019 2065

Maskin Eric and Jean Tirole "A Theory of Dynamic Oligopoly I & II " *Econometrica* 56 no 3 1988 : 549 600

Mazzeo Michael "Competitive Outcomes in Product Differentiated Oligopoly " *Review of Economics and Statsitsics* 84 no 4 2002 : 719 728

McFadden Daniel "Econometric Models of Probabilistic Choice " Edited by C F Manski and D McFadden *Structural Analysis of Discrete Data with Econometric Applications* MIT Press 1981: 198 272

McFadden Daniel "Modeling the Choice of Residential Location " Edited by A Karlqvist L Lundqvist F Snickars and J Weibull *Spatial Interaction Theory and Planning Models* 1978: 75 96

Mundlak Yair "On the Pooling of Time Series and Cross Section Data " *Econometrica* 46 1978 : 69 85

Nelson R A "Values of U S Class 1 Railroads " *Association of American Railroads, Consultant Report* 1975

Nevo Aviv "A Practitioner's Guide to Estimation of Random Coefficients Logit Models of Demand " *Journal of Economics and Management Strategy* 9 no 4 2000a : 513 548

Nevo Aviv "Measuring Market Power in the Ready to Eat Cereal Industry " *Econometrica* 69 2001 : 307 342

Nevo Aviv "Mergers with Differentiated Products: the Case of the Ready to Eat Cereal Industry " *RAND Journal of Economics* 31 2000b : 395 421

Olley G Steven and Ariel Pakes "The Dynamics of Productivity in the Telecommunications Equipment Industry " *Econometrica* 64 no 6 1996 : 1263 1297

Pakes Ariel "Patents as Options: Some Estimates of the Value of Holding European Patent Stocks " *Econometrica* 54 no 4 1986 : 755 794

Pesendorfer Martin "Horizontal Mergers in the Paper Industry " *RAND Journal of Economics* 34 no 3 2003 : 495 515

Pesendorfer Martin and Philipp Schmidt Dengler "Asymptotic Least Squares Estimators for Dynamic Games " *Review of Economic Studies* 75 2008 : 901 928

Pesendorfer Martin and Philipp Schmidt Dengler "Identification and Estimation of Dynamic Games " May 2003

Pesendorfer Martin and Philipp Schmidt Dengler "Sequential Estimation of Dynamic Discrete Games: A Comment " June 2009

Rust John "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher " *Econometrica* 55 no 5 1987 : 999 1033

Rust John "Structural Estimation of Markov Decision Processes " Edited by Robert F Engle and Daniel L McFadden *Handbook of Econometrics* 4 1994 : 3081 3143

Ryan Stephen "The Costs of Environmental Regulation in a Concentrated Industry " 2006

Semykina Anastasia and Jeffrey M Wooldridge "Estimating Panel Data Models in the Presence of Endogeneity and Selection: Theory and Application " June 2005

Surface Transportation Board "Study of Railroad Rates: 1985 2007 " 2009

Sweeting Andrew "Dynamic Product Repositioning in Differentiated Product Markets: The Case of Format Switching in the Commercial Radio Industry " *NBER Working Paper* 2007

Train Kenneth E *Discrete Choice Methods with Simulation.* Cambridge University Press 2002

U S Department of Commerce "Statistical Abstract of the U S " Annual Volumes 1978 2006

Vellturo Christopher Allan "The Deregulation of the U S Rail Industry: Efficiency and Equity in Attaining Rail Viability " PhD Thesis MIT 1989

Waters William G "Evolution of Railroad Economics " *Research in Transportation Economics* 20 2007 : 11 67

Whinston Michael D "Antitrust Policy Toward Horizontal Mergers " *Handbook of Industrial Organization* 3 2007 : 2369 2440

Williamson Oliver "Economies as an Antitrust Defense: Welfare Tradeoffs " *American Economic Review* 58 1968 : 18 36

Wilson Wesley W "Cost Savings and Productivity in the Railroad Industry " *Journal of Regulatory Economics* 11 1997 : 21 44

Wooldridge Jeffrey M *Econometric Analysis of Cross-Section and Panel Data.* MIT Press 2003

Wooldridge Jeffrey M "Selection Corrections for Panel Data Models under Conditional Mean Independence Assumptions " *Journal of Econometrics* 68 1995 : 115 132

APPENDIX

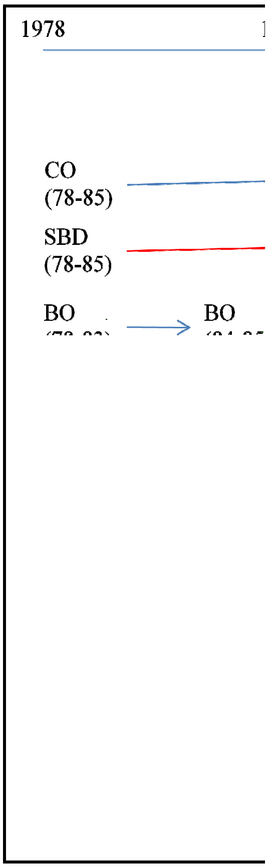

1978
CO
(78-85)
SBD
(78-85)
BO
BO

Figure 3. Network configuration of the US rail freight railroads in 2009

(Source: http://www.cn.ca/en/cn-and-class-1-railroads-flash.htm)

Figure 4. Evolution of the Class 1railroads market shares in the US freight national market

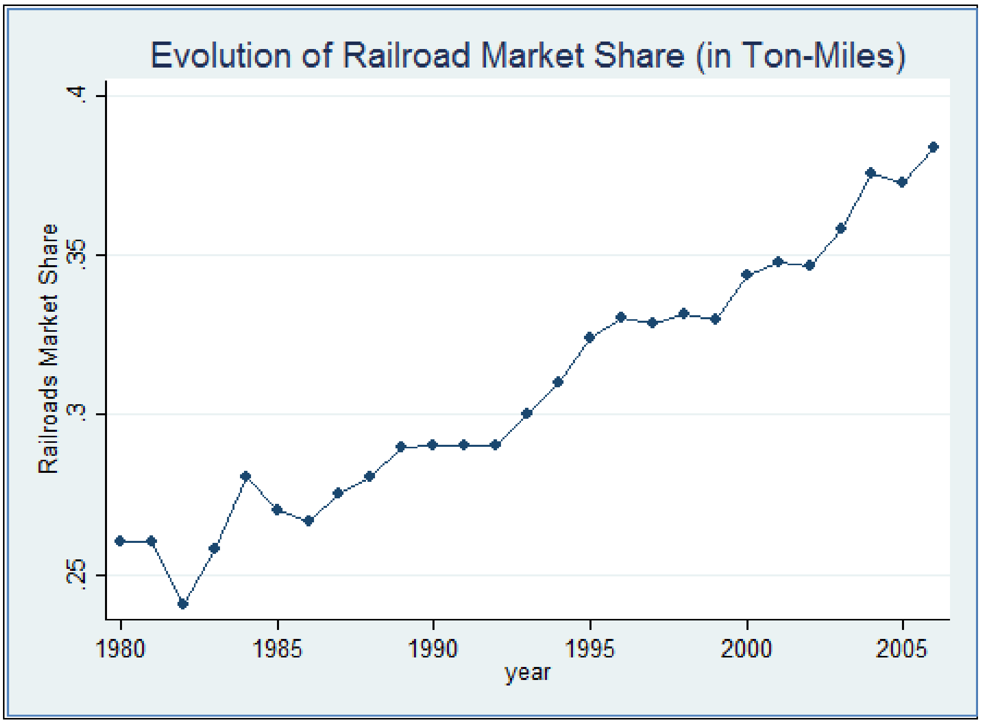


Figure 5. Industry price index (unit of measure: ton-miles, in real 1982$ = 100)

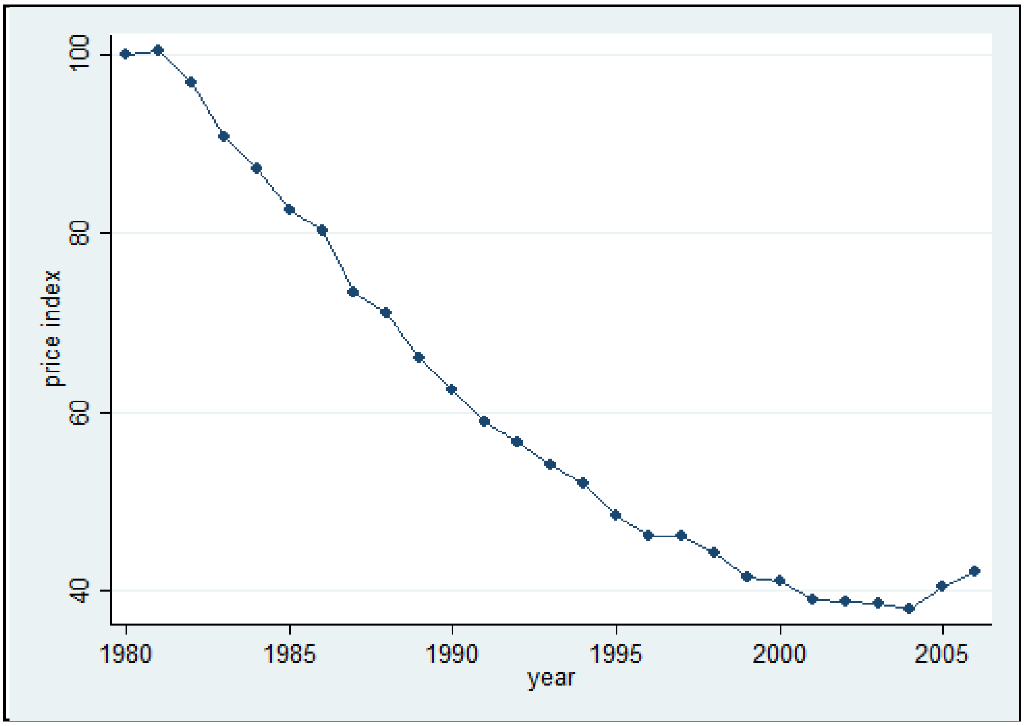
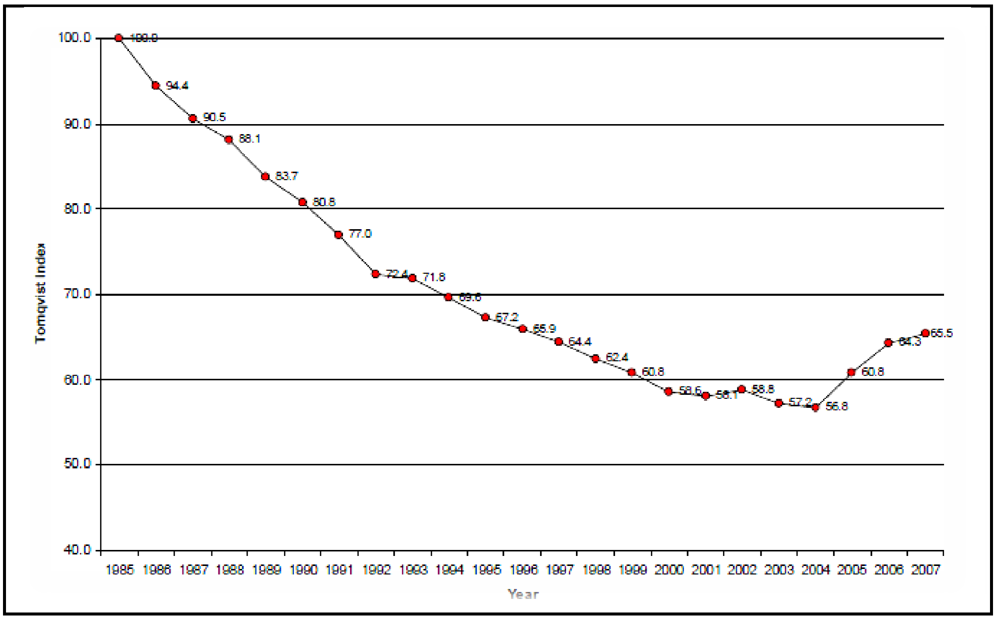


Figure 6. Rail rate index (1985 to 2007). Real revenue per ton-miles (1985 = 100)



(Source: Surface Transportation Board)

Figure 7. Evolution of marginal costs over time (1982 = 100)

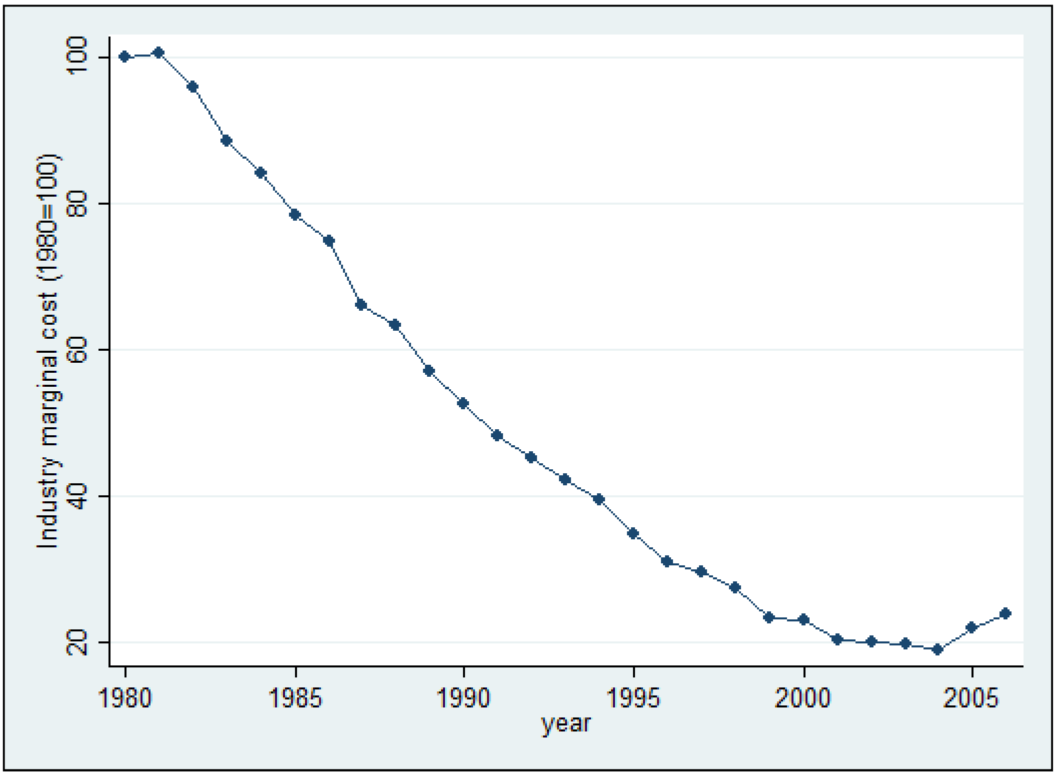


Figure 8. Evolution of markups over time (1980=100)

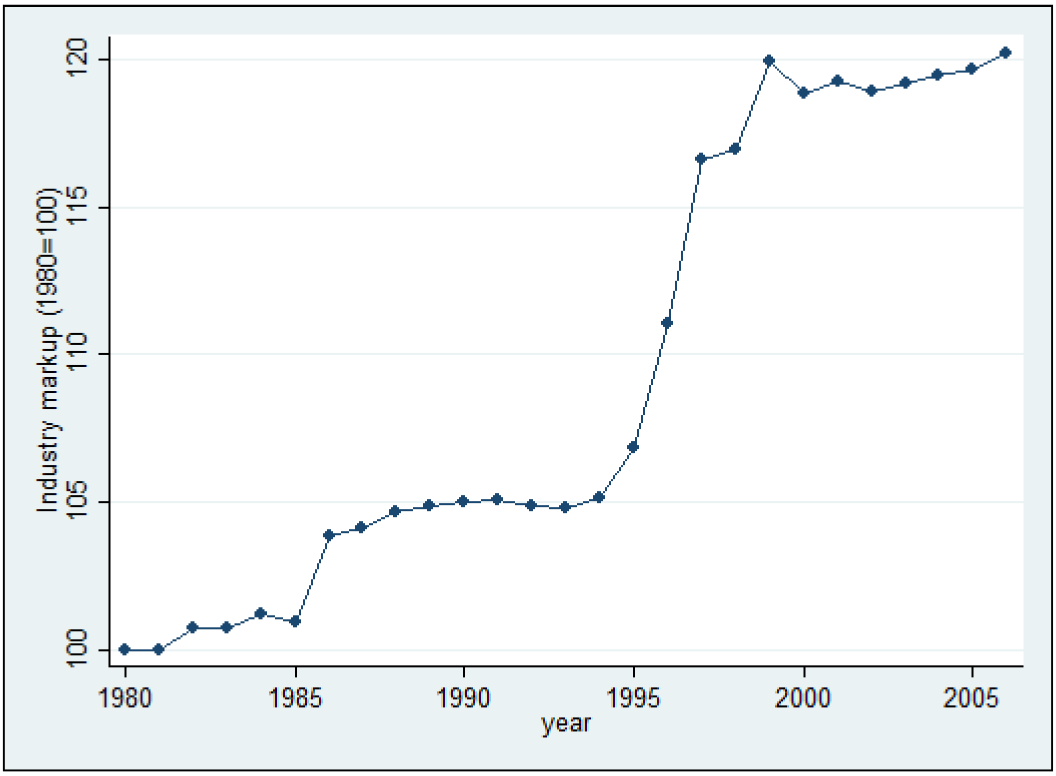


Figure 9. Evolution of consumers' surplus over time (1982=100)

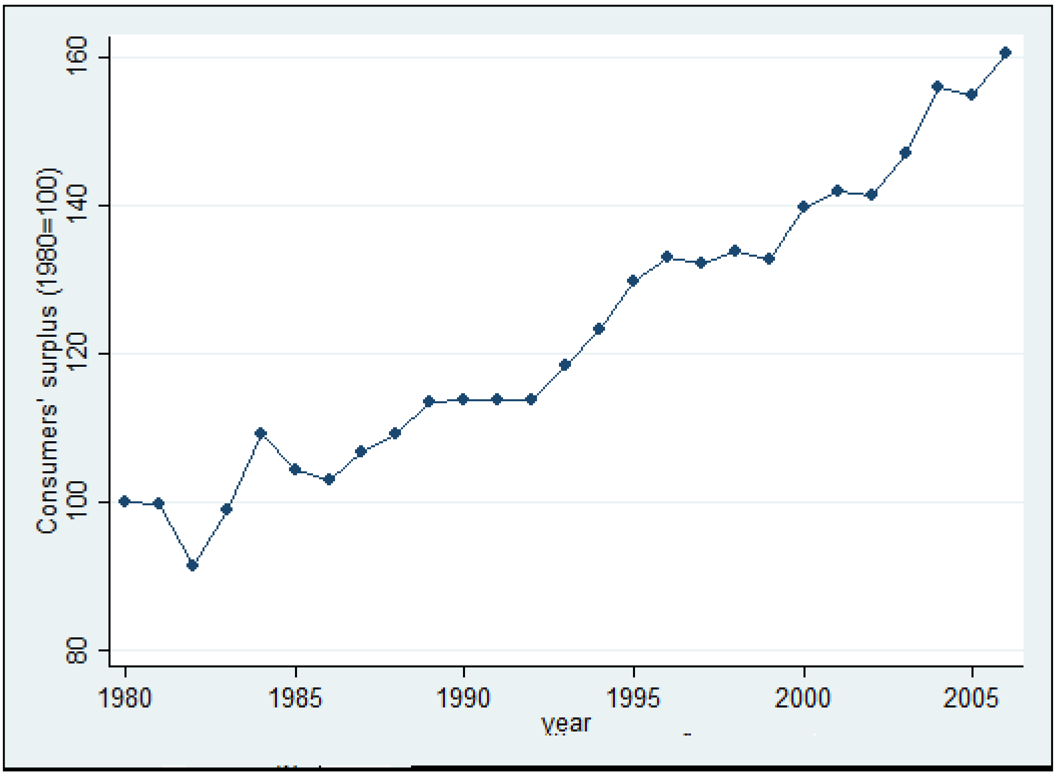


Figure 10. Evolution of the industry efficiency index over time

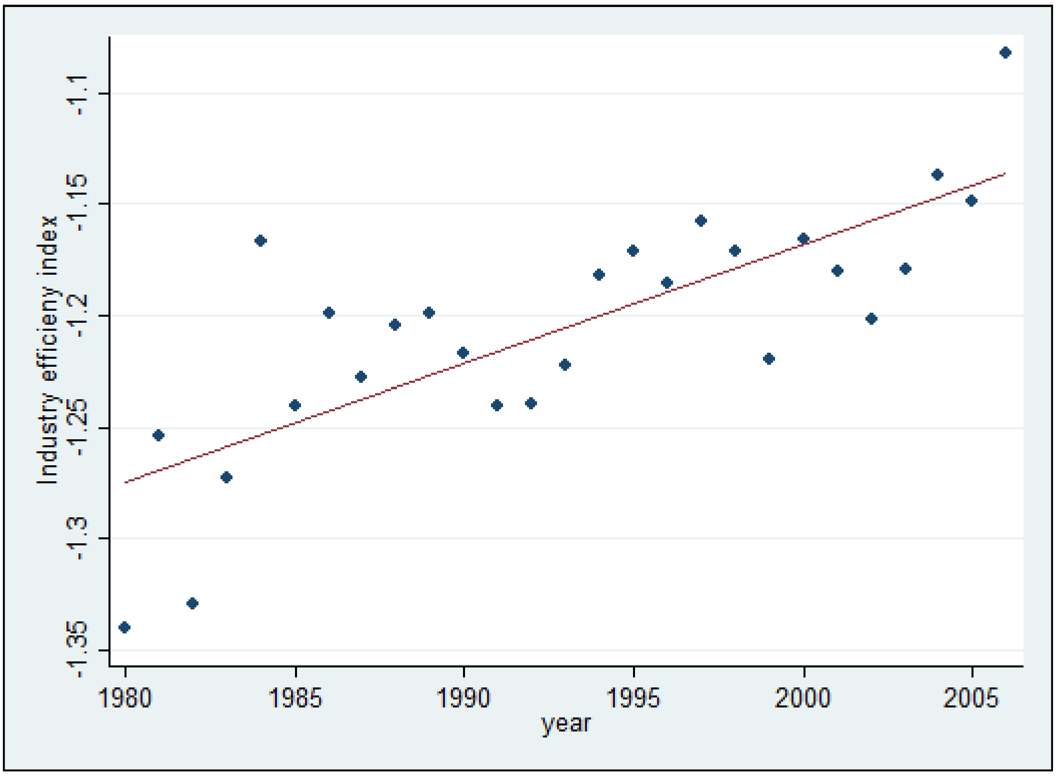


Table 1. Railroad firms in the Western area

Railroad	Years in data	Abbrevation used in Figure 1 and Figure 2
Atchison Topeka & Santa Fe ATSF	1978 1995	ATSF into with BN in 1995
Burlington Northern BN Burlington Northern Sante Fe BNSF	1978 2006	BN BNSF
Canadian National Grand Trunk Corporation CNGT	2002 2006	CNGT it incorporates all US activities of Canadian National Railroad which included GTW activities
Chicago & Northwestern CNW	1978 1994	CNW into UP in 1994
Colorado and Southern CS	1978 1981	CS into BN in 1981
Denver Rio Grande & Western DRGW	1978 1993	DRGW into SP in 1993
Detroit Toledo & Ironton DTI	1978 1983	DTI into GTW in 1983
Forth Worth and Denver FWD	1978 1981	FWD into BN in 1981
Grand Trunk & Western GTW	1978 2001	GTW
Illinois Central Gulf IC	1978 1998	IC into GTW in 1998
Kansas City Southern KCS	1978 2006	KCS
Milwaukee Road MILW	1978 1984	MILW into SOO in 1984
Missouri Kansas Texas MKT	1978 1987	MKT into UP in 1987
Missouri Pacific MP	1978 1985	MP into UP in 1985
Saint Louis and San Francisco SLSF	1978 1979	SLSF into BN in 1979
Saint Louis Southwestern SSW	1978 1989	SSW into SP in 1989
SOO Line SOO	1978 2006	SOO
Southern Pacific SP	1978 1996	SP into UP in 1996
Union Pacific UP Union Pacific Southern Pacific UPSP	1978 2006	UP UPSP
Western Pacific WP	1978 1985	WP into UP in 1985

Table 2.Railroad firms in the Eastern area

Railroad	Years in data	Abbrevation used in Figure 1 and Figure 2
Baltimore & Ohio BO	1978 1985	BO into CSX in 1985
Chesapeake & Ohio CO	1978 1985	CO into CSX in 1985
Consolidated Rail Corp CR	1978 1998	CR splitted between CSX and NS in 1999
CSX Transportation CSX	1986 2006	CSX
Norfolk Southern NS	1986 2006	NS
Norfolk & Western NW	1978 1985	NW into NS in 1985
Seaboard System Railroad SBD	1978 1985	SBD into CSX in 1985
Southern Railway System SOU	1978 1985	SOU into NS in 1985
Western Maryland WM	1978 1983	WM into BO in 1983

Table 3. Exogenous variables constructed on a firm specific basis

Variable	Source
Coal Consumption	US Department of Energy Energy Information Administration State Energy data System
Coal Production	US Department of Energy Energy Information Administration State Energy Data System
New Automobile Registration	US Department of Transportation Federal Highway Administration
State Population	US Department of Commerce Statistical Abstract of The United States and United States Department of Agriculture
Oil Price	US Department of Energy Energy Information Administration Domestic Crude Oil First Purchase Prices by Area
Farm Income	US Department of Agriculture Economic Research Service
Value of Shipment	US Department of Commerce Annual Survey of Manufacturers 2000 2006 available only for the period 1997 2006

Table 4. Descriptive statistics on US class1 railroad data

Variable	Unit	Obs	Mean	Std Dev	Min	Max
Price	$1982	353	0260265	0114441	0103483	0853767
Output	Ton Miles	353	9 21e+07	1 26e+08	1285901	6 42e+08
$K_{j\,t}$	$1982 000	353	3289 989	2842 068	141 6636	11715 29
$k_{j\,t} = \ln K_{j\,t}$	$1982	353	7 604252	1 111563	4 953455	9 368649
Coal Consumption	Thousand Short Tons 000	353	265 736	169 3512	11 98112	681 3316
ROAD	1000 Miles	303	10 34329	9 131503	0 527	35 208
HAUL	100 Miles	303	5 992068	8 505216	1 75	142 33

Table 5. Demand estimates

	Case 1	Case 2	Case 3	Case 4
attrition serial correlation	No No	Yes No	No Yes	Yes Yes
price α)	38 9059 ** 15 4278)	44 4430 *** 16 2711)	39 5622 ** 15 2709)	56 770 ** 23 0159)
σ	0 8582 *** 0 2339)	0 7232 ** 0 3448)	0 6994 ** 0 3409)	0 6639 ** 0 2872)
serial corr term: γ	NA	NA	Not estimated see eq 30))	0 1900 0 2252)
correction term: ρ	NA	0 6161 ** 0 2568)	NA	0 7450 *** 0 2831)
$k_{j\ t}$	0 0645 0 0889)	0 0910 0 2435)	0 0746 0 2291)	0 1446 0 2023)
CCON	0 0009 0 0009)	0 0007 0 0009)	0 0008 0 0009)	0 0008 0 0009)
temps	0 0858 *** 0 0313)	0 0694 * 0 0361)	0 0772 ** 0 0364)	0 0860 * 0 0506)
temps2	0 0023 *** 0 0006)	0 0018 ** 0 0007)	0 0022 *** 0 0007)	0 0021 ** 0 0009)
N deg freedom sargan p value	303 9 31 4801 0 0002	303 8 14 5278 0 0690	303 8 28 8672 0 0003	303 8 5 0194 0 7555
Standard errors in parentheses * p<0 10 ** p<0 05 *** p<0 01				

Note: The number of observations decreases from 353 to 303 due to the lag in the correction term and due to the use of lag instruments Standard errors in parenthesis) are computed by cluster bootstrap

Table 6. The staying/exit policy function

Proba of staying	Coeff
J_t	3636421 * 1953)
$v_{j\,t}$	2 205831 ** 9586)
$\sum_{l \neq j}^{J_t} v_{l\,t}$	1590722 2211)
$\xi_{j\,t}$	5868612 5640)
$\sum_{l \neq j}^{J_t} \xi_{l\,t}$	1499776 1122)
Standard errors in parentheses * p<0 10 ** p<0 05	

Note: The estimates of the quadratic time trend and the constant term are not reported Number of observations: 353 Standard errors in parenthesis) are computed by cluster bootstrap

Table 7. Descriptive statistics for mark-ups and marginal costs

	Obs	Mean	Std Dev	Min	Max
mark ups $1982)	353	0063271	0005249	0059278	0082942
marginal costs $1982)	353	0196994	0116824	0022647	0794013

Table 8. Descriptive statistics for the consumer surplus

	Obs	Mean	Std Dev	Min	Max
Consumer welfare in billion $1982)	27	0065337	0010357	0048455	0085269

Table 9. Impact of prices and efficiency on consumer welfare

	Coeff
Industry efficiency index	0055997 0004263)
Industry price index	2681975 026683)
constant	0223209 0013912)
time	0004874 0000524)
time squared	0000127 1 01e 06)
All the coefficients are significant at 1% R squared = 0 99	