Evaluation of a Taxi Sector Reform: a Real Options Approach*

Meritxell Albertí
Ercros

Angel León†
Universidad de Alicante

Gerard Llobet
CEMFI

October 19, 2003

Abstract

This paper applies the real options methodology to estimate the impact of the reestructuring measures in the taxi sector in Barcelona in 1995. These measures included among other things the establishment of a fixed price for the license as a way to reduce the uncertainty on the future performance of the sector.

We construct a model that takes into account the uncertainty in income as well as costs of the taxi driver using both individual and aggregate data. We take into account the option to sell the license at each point in time.

We show that such reform can increase the net surplus value for taxi drivers respect to a liberalized market in this sector. We compute the surplus value as a result of this reform to be around €80.000. This value is shown to be robust to an exhaustive sensitivity analysis.

Keywords: license price, net surplus value, real options, Monte Carlo simulation, continuous-time model, model implementation.

JEL Classifications: C15, C61, G13, G31 and L51.

*This paper is based on the Master project by Albertí at CEMFI (2002) under the supervision of the other two authors. Contact Author: Angel León Valle, Dpto. Economía Financiera, Universidad de Alicante, San Vicente del Raspeig, 03080 Alicante, Spain. E-mail: aleon@ua.es.

†Angel León acknowledges the financial support provided by Ministerio de Ciencia y Tecnología grant BEC2002-03797 and the Generalitat Valenciana grant CTIDIA/2002/103. The contents of this paper are the sole responsability of the authors.

1 Introduction

The taxi industry has always been the subject of some controversy in the literature of transportation economics. In this sector, an almost unique combination of regulations coexist. On the one hand, prices are regulated among all producers (taxicabs) while entry is also restricted through usually transferable licenses (sometimes denoted as medallion after the shape that this identification has in New York). Quality is also part of this regulation. Some authors, such as Cairns and Liston-Heyes (1996) argue that without such regulation this market would lack a market equilibrium due to hold-up problems.

In most developed countries the history of the sector has been similar. While the industry started as a deregulated market, the low entry costs induced waves of new taxi drivers in the periods of recession, in particular during the 1930s. As a result, prices fell rapidly and working conditions became particularly rough. Organized lobbies succeeded in putting the industry under public control, allowing to regulate fares and the number of licenses that were granted.

Nowadays, in most cities, fares are usually set through bargaining between taxi unions and the municipalities. They are subject to adjustments in order to reflect changes in costs. The number of licenses is usually related to population and it is updated when population increases. Finally, taxi drivers are evaluated according to some standards that include knowledge of the city they serve.

Cities usually differ in their treatment of licenses. For example, some cities have experimented with the complete liberalization of the sector, especially in the west of the United States, during the 1980s with mixed results. The outcome in these cities has been a substantial entry, together with a (surprising) considerable increase in prices.[1]

Licenses are in general transferable, and entry is substantially limited. In some cities, licenses are perpetual and they can be traded in quite an active market. Meanwhile, in others, the duration of the license is limited in time. Until recently, Barcelona belonged to a third group of cities where licenses were also traded but at a fixed price decided by the municipality (denoted as the "Borsa de Llicències" or "License Bag"). To avoid side-payments these licenses were bought by the municipality, that would later assign them to new candidates. This system was introduced in Barcelona in 1995 as a response to the crisis in this sector. The restructuring process undertaken by the city was combined by several other measures being the most important the reduction in the number of licenses. This system was in place until 2002 where the license price was again liberalized once the sector recovered.

The regulated price of the license was set to €36,000 in 1995. This amount was partially reflecting the market price for the license around that time. The price was barely updated in the following years and it was around €36,500 in 2002. This price is considerably lower than the €60,000 to €80,000 paid for licenses in other cities in Spain, such as Valencia and Madrid.[2] The failure in updating the license price was the reason that leaded to the liberalization of this

[1] Teal and Berglund (1987) estimate increases in the number of taxis ranging from 18% in Kansas City to 127% in San Diego. In the mean time, prices rose in San Diego by 26.2% with similar performance in other cities.

[2] See *El Periódico de Cataluña* (12/4/2001).

sector after pressure from taxi drivers in 2002.[3]

This paper estimates how much the fixed price system for the licenses (together with the other measures) would have benefited the taxi drivers compared to the liberalized price common in most other cities when it was first applied in 1995. To do so, we take into account that owning a license does not only entitle to a future stream of cash-flows – before retirement – derived from the work of a taxi driver but also to the right to resell the license at any time for a preespecified price and exert another profession. In contrast, in a free license market where the price is set in equilibrium, all aggregate information is already incorporated in the price of the license which means that selling the license can only protect the taxi driver against idiosyncratic uncertainty. In other words, the price of the license essentially reflects the present value of the cab driver's profits in the future periods. Hence, the net surplus[4] of a taxi driver in a market where the price of the license is not regulated, should be approximately zero. Summing up, the aim of this paper consists of computing the cab driver's net surplus under a fixed price framework for licenses settled in Barcelona city in 1995.

A regulated license price, however, insures the taxi driver against aggregate shocks in the present value of profits of his profession. Therefore, how much the option to sell the license at a fixed price adds to profits is partially due to the increase in utility of this mechanism. The results we obtain show that the net surplus of a typical taxi driver in the fixed license mechanism would have amounted to €80,000. This number should be interpreted as how much a taxi driver would have been willing to pay to participate in the sector after the restructuring process initiated in 1995 when he expected to work in the sector for around 30 years.

So why did the mechanism break down in Barcelona? The most likely reason is the time inconsistency in the decision of taxi drivers. Although the original *Borsa de Llicències* benefited entrants by lowering the price of the license they had to pay, once they were insiders they became interested in liberalizing the sector as a way to obtain a higher price once they sold the license.

The environment under which this study is conducted is especially useful. The authority guaranteed a future price for the license, reducing uncertainty regarding its resale. We also have estimates of costs and income of a typical taxi driver from a study conducted by the Metropolitan Institute of Taxi in Barcelona (IMET) in 1995 in order to assess on the changes in the sector. In this study, we will also use other sources of data so as to estimate, for example, the evolution of fuel costs, the growth rate of real gross domestic product and so on. All these estimates will be necessary to obtain the net surplus implied from our theoretical model.

We model the decision of the taxi driver either to stay in the market or sell the license as an *American put option*. In other words, when the taxi driver buys the license, he acquires the rights not only to a future stream of income (and cost) but also the possibility to leave this job and sell this right for a preespecified price at any moment.

The value of the taxi driver project, that is the present value of the stream of profits once the abandonment flexibility is considered, can be estimated using several methodologies. In order to approximate the project value, we rely on the algorithm presented by Longstaff

[3] See *El Periódico de Cataluña* (3/13/2002).

[4] It is defined as the present value of the taxi-driver's profits once we deduct the relevant opportunity costs and the price of the license. Later, it will be explained in more detailed.

and Schwartz (2001) that consists of a simulation procedure. This algorithm compares the value of exercising the option to abandon implicit in the project with the expected value of continuation at different points of time in each of the simulated trajectories. The key feature to this approach is to estimate, using least squares (LS) approach, the conditional expected payoff to continuation. In recent years this methodology has been applied to several problems. For example, Schwartz and Moon (2000) used it to value internet firms and Schwartz (2001) to value patents and R&D projects, together with the optimal policy for the management of innovation. This algorithm is easy to implement and it is a very good alternative against the well-known finite difference method which becomes impractical for solving partial differential equations when there are multiple factors as our model exhibits.

In our case, we use the above methodology applied to a complex multifactor model where the state variables are mainly stock, together with flow ones. Their dynamics are driven, for example, by jump-diffusion processes, which approximate phenomena such as fare negotiation processes, and so forth. Our experiment allow us to value the cost of aggregate shocks in service sector, compared to the case where these shocks were not present. We also make a thorough sensitivity analysis to verify how robust are our results. In this case, we emphasize one of our contributions. We provide a systematic methodology to study how each parameter, or a subset of them, contribute to the final net surplus value and we do it for different sample sizes to either calibrate or estimate, depending on the equation in the model, some parameters.

The rest of the paper proceeds as follows. In section 2 we formulate the model and the dynamic equations that describe the process for costs and revenues of the taxi driver. Section 3 presents the solution procedure and describes the Longstaff and Schwartz Monte Carlo algorithm (LSM). Section 4 describes the sources of our data, how we calibrate or estimate the different parameters of the model and presents the main results. Section 5 performs sensitivity analysis regarding the main variables of the model and section 6 presents some extensions. Section 7 concludes.

2 Continuous-time model

The model studies the decision to become a taxi driver and in particular the willingness to pay for a license. Licenses have a positive value because they are in limited supply, controlled by the public authority. The driver considers at each point in time the present value of income and costs, taking into account the fact that the license could be sold, and he could work in another sector earning an alternative wage. To the extent that this option might be executed, the value of a license would be substantially different from the present value of cash flows of the driver.

We assume that the taxi driver can work for at most T years, after which the license must be sold due to retirement. We are also considering that the driver can abandon the project only once, moment after which he can not reenter the market.

At the initial moment, the agent decides whether to become a taxi driver. In this case, besides the cost of the license, a car needs to be bought. After that, the driver receives a flow of revenues in each period t, $R(t)$ and incurs in variable costs $C(t)$ and fixed costs $F(t)$. So, the cash flow available to the taxi driver at the end of period t, denoted as $CF(t)$, is:

$$CF(t) = R(t) - C(t) - F(t). \tag{1}$$

Costs are mainly driven by the price of fuel, which is the largest component of these costs. Fixed costs are essentially related to the cost of the car and the regular maintenance.[5]

Revenues are a function of both the number of trips that a taxi driver makes and the prices charged. We assume that the number of trips is a function of the business cycle, while prices are adjusted according to the negotiation between the taxi union and the municipality. New prices depend on changes in costs.

We define next the processes we assume for costs and revenues in detail.

2.1 Cost equation

Total costs at each point in time are generated by three components. The first are *variable costs*, depending on both the distance in kilometers and the car type. Its main component is the cost of diesel oil,[6] together with other variable costs such as lubricants, tires, maintenance and repair expenses. The second component are *fixed costs* that contain mainly income and value added taxes, car insurance and social contributions. Other fixed costs are administration, parking, vigilance and the cost of buying a car in case of an independent driver. Finally there are *financial costs* related to financing the car purchase.

We will only consider as a stochastic variable the item corresponding to the fuel cost due to the dynamics of the fuel price. So, we will lump together all the variable costs other than fuel with the fixed and the financial costs. We denote them by $F(t)$ that is assumed to be deterministic with a growth rate equal to the inflation rate, i.e.

$$dF(t) = \pi(t) F(t) dt \tag{2}$$

where $\pi(t)$ denotes the inflation rate per year that is assumed to be non-stochastic to simplify the model. Meanwhile, the evolution of the only source of uncertainty, the fuel bill, is assumed to be given by the one-factor model for commodity spot prices proposed by Schwartz (1997) which follows the stochastic process

$$dS(t) = k_s(\mu_s - \ln S(t)) S(t) dt + \sigma_s S(t) dW_s(t). \tag{3}$$

Let $X(t) \equiv \ln S(t)$ and after applying Ito's lemma, the log price follows an Ornstein-Uhlenbeck process defined by

$$dX(t) = k_s(\alpha - X(t)) dt + \sigma_s dW_s(t) \tag{4}$$

where

$$\alpha = \mu_s - \frac{\sigma_s^2}{2k_s}.$$

[5]See section 4.1.1 for more details.

[6]In 1995, 87% of total cabs used diesel oil, 10% butane gas and the rest gasoline.

This process assumes that the log price mean reverts to the long-term level α at a speed given by the mean reversion rate k_s which is considered to be strictly positive, and with a long run mean for the spot price $\overline{S} = \exp(\mu_s)$. As a result, $S(t)$ follows a lognormal distribution. Notice that the volatility of the logarithm of $S(t)$ is σ_s.

Let $\gamma_\Delta(t)$ be the annual return or growth rate of $S(t)$ for a time interval of length Δ, defined as

$$\gamma_\Delta(t) = \frac{1}{\Delta}\int_{t-\Delta}^{t} dX(u) = \frac{X(t) - X(t-\Delta)}{\Delta}. \tag{5}$$

Then the instantaneous return or growth rate of $S(t)$, denoted as $\gamma(t)$, is given by

$$\gamma(t) = \lim_{\Delta\to 0}\gamma_\Delta(t) = \frac{1}{S(t)}\frac{\partial S(t)}{\partial t}.$$

Let $C(t)$ denote the cost of gas oil (variable cost) from operating a taxicab. We impose that its dynamics are given by

$$\frac{dC(t)}{C(t)} = \gamma(t)\,dt + \sigma_s dW_c(t). \tag{6}$$

The growth rate and volatility of $dC(t)/C(t)$ are governed by $\gamma(t)$ and σ_s. Denote as $dW_c(t)$ another standard Wiener increment defined as

$$dW_c(t) = \rho_{sc} dW_s(t) + \sqrt{1-\rho_{sc}^2}\, dW_a(t). \tag{7}$$

We will assume that $dW_c(t)$ is highly correlated with respect to $dW_s(t)$, so that ρ_{sc} is large, since the uncertainty is mainly driven by the evolution of gas oil prices. There is, however, another component of uncertainty represented by the Wiener process $dW_a(t)$ that captures the consumption in gas oil depending on other variables such as the car-type, the distance in kilometers or the number of hours worked by the particular driver.

2.2 Revenue equation

The revenue from operating a taxicab, $R(t)$, is accrued mainly from the fare multiplied by the number of passengers "service units" supplied.[7] Between changes in tariffs, the behavior of revenues depends on a demand that is mainly determined by the business cycle.[8] Trivially, changes in revenues can be decomposed as changes in quantities, $q(t)$, and changes in prices, $p(t)$,

$$\frac{dR(t)}{R(t)} = \frac{dq(t)}{q(t)} + \frac{dp(t)}{p(t)}. \tag{8}$$

[7] Additional income for advertising in cab is not considered in our model.

[8] This evidence is supported by studies carried out by IMET. They also show more than half of all trips are done for business reasons.

2.2.1 Changes in quantities

In particular, we assume that the dynamics of the service units provided are generated by the following stochastic differential equation (henceforth, sde):

$$\frac{dq(t)}{q(t)} = \theta_q(t)\,dt + \sigma_q dW_q(t) \tag{9}$$

where $dW_q(t)$ is a standard Wiener increment with constant volatility σ_q. The instantaneous drift of this process is denoted by $\theta_q(t)$. This term captures the stochastic demand realizations conditioned to fixed fares. Because demand is extremely inelastic, we approximate these changes as independent of prices. Therefore, since quantities are very procyclical, the drift $\theta_q(t)$ or expected growth rate in quantities is approximated by the growth rate or return of the real gross domestic product (RGDP).[9] This drift evolves according to the following mean reverting process:

$$d\theta_q(t) = k_\theta\left(\overline{\theta} - \theta_q(t)\right)dt + \sigma_\theta dW_\theta(t) \tag{10}$$

where $\overline{\theta}$ denotes the long-term average drift, k_θ is the mean reversion rate which is assumed to be strictly positive such that $\ln(2)/k_\theta$ can be interpreted as the "half-life" of the deviations from $\overline{\theta}$, σ_θ is the volatility of the drift and $dW_\theta(t)$ is another standard Wiener increment that is independent of $dW_q(t)$. In principle we assume that $dW_\theta(t)$ is also correlated with $dW_c(t)$ so that,

$$dW_c(t)\,dW_\theta(t) = \rho_{c\theta}dt. \tag{11}$$

This relationship assumes that higher oil prices might have negative effects on growth and therefore reduce the rate of increase in passenger income. To the extent that the growth rate is affected by oil prices, this relationship will be negative and so $\rho_{c\theta} \leq 0$.

Notice that we are assuming that the quantities sold do not depend on changes in population. The reason is that the number of licenses is often revised to accommodate increases in the number of inhabitants of the city.

2.2.2 Changes in prices

Changes in prices are assumed to be "rare" and stochastic in nature, which are the result of negotiation between taxi unions and the municipalities. In particular, changes in prices are formulated as the following compound Poisson process:

$$\frac{dp(t)}{p(t)} = \varphi(t)\,dN_p(t) \tag{12}$$

where $dN_p(t)$ is the Poisson process with intensity $\lambda_p dt$ and $\varphi(t)$ is the growth rate of the fare. The process $dN_p(t)$ is assumed to be independent of both $dW_q(t)$ and $dW_\theta(t)$. Since the rule that governs exactly the change in fares is unknown, we will suppose that when a change in

[9] We relax this assumption in Section 6 without significant results.

fares occurs its revision could be a function of the evolution of the GDP deflator, henceforth GDPD. We will identify $D(t)$ as the GDPD whose dynamics are governed by the following geometric Brownian motion

$$\frac{dD(t)}{D(t)} = \mu_D dt + \sigma_D dW_D(t). \tag{13}$$

We will allow for correlation between the growth rate of real GDP and the GDP deflator to exist, so that $dW_\theta(t)\, dW_D(t) = \rho_{\theta D}$.

The common practice when changing the fares is to adjust its increase by the change in prices after the last revision occurred. Analytically, this rule is represented as follows: if t' denotes the last time a change in fare occurred and the fare is revised again in t, obviously $t' < t$, then the variation is measured using the growth rate of GDPD for the period. Applying Ito's lemma on (13), the growth rate $\varphi(t) \equiv \ln(D(t)/D(t'))$ is obtained as

$$\ln(D(t)/D(t')) = \beta_D \delta t + \sigma_D \sqrt{\delta t}\xi_D(t) \tag{14}$$

where $\beta_D \equiv \mu_D - 0.5\sigma_D^2$, $\delta t \equiv t - t'$ and $\xi_D(t) \sim iid\ N(0,1)$. Note that $\varphi(t)$ represents the log of a jump size. Therefore, the price after the revision becomes

$$p(t) = \exp(\varphi(t))\, p(t').$$

that corresponds to the solution at time t of the geometric Brownian motion

$$\frac{dp(s)}{p(s)} = \mu_D ds + \sigma_D dW_D(s), \qquad s \in [t', t].$$

We also assume that each revision (jump size) is independent of the previous ones. In other words, let $t_1 < t_2 < \ldots < t_n$ be n dates where a fare alteration has happened, then the conditional distribution for the log of jump size verifies that

$$\varphi(t_i)\,|I_{t_i-1} \backsim id\ N\left(\varphi(t_{i-1}) + \beta_D \delta t_i,\ \sigma_D^2 \delta t_i\right)$$

where I_{t_i-1} is the information set evaluated at t_{i-1} and $\delta t_i \equiv t_i - t_{i-1}$.

Summing up, given the previous equations we will refer to changes in revenues according to equation (8) as described exactly by

$$\frac{dR(t)}{R(t)} = \underbrace{\theta_q(t)\, dt + \sigma_q dW_q(t)}_{\frac{dq(t)}{q(t)}} + \underbrace{\varphi(t)\, dN_p(t)}_{\frac{dp(t)}{p(t)}}. \tag{15}$$

2.3 Risk-neutral measure

Notice that equations (4), (10) and (15) are under the real measure. With respect to the process for $R(t)$ we will assume that the jump risk is not systematic and so it is unpriced by the market. However, the uncertainty from $W_q(t)$ has a risk premium ϕ_q, assumed to be constant. Using standard techniques, the corresponding risk-neutral measure for equation (15) can be obtained by

$$\frac{dR(t)}{R(t)} = \left[\theta_q(t) - \phi_q\right] dt + \sigma_q dW_q^*(t) + \varphi(t)\, dN_p(t) \tag{16}$$

where $dW_q^*(t)$ is the Wiener increment under the risk-neutral measure so that $dW_q^*(t) = dW_q(t) + \left(\phi_q/\sigma_q\right) dt$. For simplicity, we suppose that for equation (10) the true and risk-adjusted processes are the same, so that the risk premium is zero. The dynamics of the Ornstein-Uhlenbeck process defined in (4) under the equivalent martingale measure can be rewritten as

$$dX(t) = k_s(\alpha^* - X(t))\, dt + \sigma_s dW_s^*(t). \tag{17}$$

Notice that we have defined $\alpha^* = \alpha - \phi_s$, where ϕ_s is the risk premium (also assumed constant) of the gas oil price risk and $dW_s^*(t) = dW_s(t) + (\phi_s/\sigma_s)\, dt$ denotes again the Wiener increment under the equivalent martingale measure. Finally, we also assume a zero risk premium for the uncertainty $dW_a(t)$, so that the risk-neutral measure $dW_c^*(t)$ is written as

$$dW_c^*(t) = \rho_{sc} dW_s^*(t) + \sqrt{1-\rho_{sc}^2}\, dW_a(t).$$

The risk neutral measure for equation (6) is

$$\frac{dC(t)}{C(t)} = (\gamma(t) - \rho_{sc}\phi_s)\, dt + \sigma_s dW_c^*(t) \tag{18}$$

where $dW_c^*(t)\, dW_\theta^*(t) = \rho_{c\theta} dt$.

2.4 Optimal stopping and license value

Let $V(y,t)$ be the project value of driving a taxi at date t. $V(y,t)$ does not only incorporate the present value of the taxi driver's stream of cash flows but it also does the flexibility of changing this job for a new one and also the income for selling the taxi driver license. We will assume that both the license price, denoted as $L(t)$, and the salary corresponding to the new job, denoted as $w(t)$, grow at the inflation rate $\pi(t)$, that is:

$$\frac{dL(t)}{L(t)} = \pi(t)\, dt, \qquad \frac{dw(t)}{w(t)} = \pi(t)\, dt. \tag{19}$$

Let y, the abbreviation of $y(t)$, denote the set of state variables that determine the value of the project. So, y contains the following variables (also represented in abbreviated form):

$$y = (q, \theta_q, C, X, F, L, w, p)'.$$

Here we consider a binary decision problem. At every instant, the taxi driver can either continue, i.e. driving the taxi, or stop and get a payoff that amounts to the income received for selling the license and the salary of the alternative job that will be earned from that moment until retirement. In this model, we will also assume that once he has decided to sell the license then he will never become a taxi driver again. Summing up, the Bellman equation[10] for our optimal stopping problem becomes

$$V(y,t) = \max \left\{ \underbrace{\Omega(y,t)}_{abandon};\ \underbrace{CF(y,t) + (1+rdt)^{-1} E^*_{y,t}\left[V(y+dy,t+dt)\right]}_{continue} \right\} \tag{20}$$

where $E^*_{y,t}[\cdot]$ denotes the conditional expected value on the set of state variables at date t and conditioned to y under the risk neutral measure, r denotes the annual risk-free rate and $\Omega(y,t)$ denotes the abandonment payoff. Our Bellman equation is subject to a boundary condition since there is a fixed time limit, denoted as T, which is the age of retirement and it will be represented as

$$V(y,T) = \Psi(y), \qquad \forall y. \tag{21}$$

Since the dynamics of the fare price $p(t)$, described in equation (12), is governed by a Poisson process, then we will assume that over a short interval of time dt, the probability of one jump happening in the fare price is $\lambda_p dt$, meanwhile $1-\lambda_p dt$ for the case of no jump. So, there are two possibilities for the change in value $dV(y,t)$, depending on whether a jump in the fare price takes us into the stopping (abandonment) region or not. If it does, then

$$dV(y,t) = \lambda_p dt\left[\Omega(y+dy,t) - V(y,t)\right] + (1-\lambda_p dt)\left[V(y+dy,t) - V(y,t)\right] \tag{22}$$

and if it remains in the continuation region, we get a similar equation but with $\Omega(y+dy,t)$ replaced by $V(y+dy,t)$. Given equations (20) and (22), we can express the continuation region as:

$$rV(y,t) = CF(t) + \frac{1}{dt}E^*_{y,t}\left[dV(y,t)\right]. \tag{23}$$

Now, rewrite y as $y=(z,p)'$ where z contain the pure diffusion processes, meanwhile p is driven by a pure compound Poisson process. Applying the multidimensional Ito's lemma with jumps[11] in (23), we obtain the following partial differential equation (henceforth, pde):

$$\begin{aligned} rV(y,t) &= CF(y,t) + \lambda_p\left[\Omega(p+dp,z,t) - V(y,t)\right] \\ &\quad + V_t(y,t) + V_z(y,t)A(z,t) + \frac{1}{2}tr\left[\Gamma(t,z)\Gamma(t,z)' V_{zz}(y,t)\right] \end{aligned} \tag{24}$$

where V_t, V_z and V_{zz} denote the obvious partial derivatives of V valued in $\mathbb{R}, 1\times\mathbb{R}^7$ and $\mathbb{R}^{7\times 7}$ respectively, $tr(\cdot)$ denotes the trace of a square matrix and z is driven by the following diffusion equation system under the risk-neutral measure:

$$dz = A(t,z)\,dt + \Gamma(t,z)\,d\mathbf{W}(t)$$

[10] For more details about dynamic optimization under stochastic continuous time, see Dixit and Pindyck (1994).

[11] See Duffie (2001).

where this system is obtained through the following equations: (2), (9), (10), (17), (18) and (19). Finally, we will rewrite the Wiener increments of these equations as functions of independent Wiener ones, denoted as $d\mathbf{W}(t)$.

For each t, assume that the pde (24) holds, that is continuation is preferred to stopping, for $y > y^*(t)$ where $y^*(t)$ is the critical value function that is unknown. This function divides the (y,t) space into two regions with continuation optimal above the curve and abandoning optimal below it. So, for the solution of the dynamic optimization program we will solve both V and y^* but for that, it will be necessary to impose the two well-known boundary conditions along $y = y^*(t)$ which are the "value-matching condition" and the "smooth-pasting condition which can be seen for a more detailed discussion in Dixit and Pindyck (1994). This procedure is difficult to implement since the pde in (24) has no closed-form solution and the large number of different factors make the use of numerical methods for pde inconvenient. Finally, another thing interfering with the above procedure is that (24) is not local to the continuation region due to the Poisson process[12]. So, we will solve the project of the value V by using a variation of the algorithm proposed by Longstaff and Schwartz (2001) that will be explained later.

Finally, once we have obtained V at the initial moment, denoted as $t = 0$, then the taxi driver's net surplus is obtained by substracting both the present value, also evaluated at $t = 0$, of the stream of the new job's wages finishing at T and the license price from the project V.

3 Model implementation

The algorithm we implement will obtain the solution numerically by discretizing the continuous time model for sufficiently small increments. For this reason, we will start with the formulae that approximate the previous processes and later, we will describe the algorithm.

3.1 Discrete approximation

Since there is no analytical solution to the continuous time model presented, it is solved using Monte Carlo simulation. The value of the project is computed using a variation of the least-squares Monte Carlo (LSM) algorithm proposed by Longstaff and Schwartz (2001) for the valuation of American options. In the simulations, we use the following discrete approximations to equations (10), (16) and (18) respectively:

$$\theta_q(t) = \overline{\theta}(1 - e^{-k_\theta \Delta}) + e^{-k_\theta \Delta}\theta_q(t-\Delta) + \sigma_\theta \sqrt{\frac{1 - e^{-2k_\theta \Delta}}{2\kappa_\theta}}\ \xi_\theta(t) \tag{25}$$

$$R(t) = R(t-\Delta)\exp\left[\theta_q(t) - (\phi_q + 0.5\sigma_q^2)\Delta + \sigma_q\sqrt{\Delta}\xi_q(t) + \xi_p(t)\, N_\Delta(t)\right] \tag{26}$$

$$C(t) = C(t-\Delta)\exp\left[\gamma_\Delta(t) - (\rho_{sc}\phi_s + 0.5\sigma_c^2)\Delta + \sigma_c\sqrt{\Delta}\ \xi_c(t)\right] \tag{27}$$

[12] See Dixit and Pindyck (1994, pp. 113).

where $\xi_\theta(t), \xi_q(t)$ and $\xi_c(t)$ are standard normal variates and the correlation between $\xi_\theta(t)$ and $\xi_c(t)$ is $\rho_{c\theta}$. The parameter Δ is the time interval and $N_\Delta(t)$ is the approximation to the continuous Poisson process $dN_p(t)$ by a Bernoulli distribution with a probability $\lambda_p\Delta$ of a jump during a time interval of length Δ.

Note that in equation (27) it is necessary to determine the dynamics for the process $\gamma_\Delta(t)$ given in (5). The next proposition describes the dynamics for $\gamma_\Delta(t)$.

Proposition 1 *The annual growth rate for a time interval Δ, $\gamma_\Delta(t)$, follows an ARMA(1,1) process defined as:*

$$\gamma_\Delta(t) = e^{-k_s\Delta}\gamma_\Delta(t-\Delta) + \epsilon(t) \tag{28}$$

where $\epsilon(t)$ is a MA(1) process[13] *with variance and first order covariance defined as*

$$\begin{aligned} E\left[\epsilon^2(t)\right] &= \frac{\sigma_s^2}{2k_s\Delta^2}\left[3 - k_s\Delta - (3 + k_s\Delta)e^{-2k_s\Delta}\right], \\ E\left[\epsilon(t)\epsilon(t-\Delta)\right] &= \frac{\sigma_s^2}{2k_s\Delta^2}\left[1 - e^{-k_s\Delta}\right]. \end{aligned}$$

Proof. See Appendix A. ■

3.2 Algorithm

A variation of the LSM algorithm proposed by Longstaff and Schwartz (2001) is implemented in order to solve the optimal policy and so, to obtain the value of a project by simulating numerous paths of the equations discretized above. In particular, denote T as the project horizon in years and let $N = T/\Delta$ be the number of periods for every path of the simulation. Equations $(25), (26)$ and (27) are used to generate M paths of $\widetilde{N} = N + 1$ periods for each of the three variables. For each path i we obtain three vectors of length $\widetilde{N} \times 1$: $R(i) \equiv [R(i, j\Delta)]$, $\theta_R(i) \equiv [\theta_R(i, j\Delta)]$ and $C(i) \equiv [C(i, j\Delta)]$ where $j = 0, 1, ..., N$. Note that for $j = 0$ we have the starting values corresponding to the year 1995 as we will see later. Finally, given the fixed cost at each point, i.e. $F(i, j\Delta)$, we obtain the value of the cash flow $CF(i, j\Delta)$ as in (1).

The algorithm searches for the optimal stopping time along each path by backwards induction. It is assumed that the option to abandon the project can only be exercised once. We find in each path i the optimal stopping time among the N possible exercise dates. The value of the project at each point in time is denoted as $V(i, j\Delta)$. Conditional on not having abandoned the project before, at the expiration date it is characterized by the following boundary condition:

$$V(i, T) = \max\left\{L(T) + w(T); \exp(-r\Delta)L(T + \Delta) + CF(i, T)\right\}.$$

[13] Notice that the MA(1) process can be represented as $\epsilon(t) = \varpi(t) + \delta\varpi(t - \Delta)$ where $\varpi(t)$ is a white noise process distributed as $N\left(0, \sigma_\varpi^2\right)$. Since δ satisfies $\rho_\epsilon\delta^2 - \delta + \rho_\epsilon = 0$ where ρ_ϵ denotes the MA(1) first order correlation, we will only consider for the discrete approximation that root which guarantees the invertibility condition for the MA(1) process.

This expression means that if the taxi driver decides to abandon in the last period, he receives both $L(T)$ for selling his license and the salary for his new job, $w(T)$, that will be only held during that period. Nevertheless, if he decides to continue then he gets the earnings from driving the taxi that amount to $CF(i,T)$ and he also receives at the end of the period $\exp(-r\Delta)\, L(T+\Delta)$ which corresponds to the price of the license for the following period discounted appropriately. For any previous date $j\Delta < T$, the present value of the project to continue at point $(i, j\Delta)$ conditional on not having abandoned the project before along path i is computed as[14]

$$V_c(i,j\Delta) = \sum\nolimits_{k=j+1}^{N} \exp(-r\,(k-j)\,\Delta)V(i,k\Delta); \qquad 0 \leq j < N \tag{29}$$

The expected value of continuation evaluated at point $(i, j\Delta)$ where $j\Delta < T$ is obtained by regressing the dependent variable in (29) onto a set of basis functions of the state variables[15] at date $j\Delta$ where the sample includes all M paths. Let $\widehat{V}_c(i,j\Delta)$ be the fitted value of the regression and let $V_a(i,j\Delta)$ denote the value of the project to abandon at date $j\Delta$ defined as

$$V_a(i,j\Delta) = L(j\Delta) + PV_w(j\Delta)$$

where $PV_w(j\Delta)$ is the present value of the alternative wage stream from the current date $j\Delta$ to period T. That is,

$$PV_w(j\Delta) = w(j\Delta)\,\frac{1-\exp(-\widetilde{r}\,(N-j+1)\,\Delta)}{1-\exp(-\widetilde{r}\Delta)}$$

where $\widetilde{r} = r - \pi$ denotes the real interest rate per year. Each $V(i,j\Delta)$ is obtained according to the following rule: if $\widehat{V}_c(i,j\Delta) + CF(i,j\Delta)$ is smaller than $V_a(i,j\Delta)$ we set

$$V(i,k\Delta) = \begin{cases} w(j\Delta) + L(j\Delta)\,; & k = j \\ w(k\Delta)\,; & j < k \leq N \end{cases}$$

For the remaining paths, where $\widehat{V}_c(i,j\Delta) + CF(i,j\Delta)$ is larger than $V_a(i,j\Delta)$, we set

$$V(i,j\Delta) = CF(i,j\Delta).$$

By rolling back in time and repeating the procedure at each date $j\Delta$, we can fill out the corresponding column of the $M \times \widetilde{N}$ matrix $V \equiv [V(i,j\Delta)]$. Notice that the optimal stopping time rule in this paper consists of obtaining for each path i the minimum $j\Delta$ such that the abandonment is better than continuing. Let $\overline{V}(j\Delta)$ represent the average of the $V(i,j\Delta)$'s at date $j\Delta$. The net surplus in the first period, $\Gamma(0)$, is obtained as follows:

$$\Gamma(0) = V(0) - PV_w(0) - L(0)$$

[14]Note that we do not take into account the cash flow at point $(i,j\Delta)$ for obtaining $V_c(i,j\Delta)$. It will be considered later for $V(i,j\Delta)$.

[15]Following Longstaff and Schwartz (2001), we consider here as independent variables: constant, R, θ_R, C, R^2, θ_R^2, C^2, $R\times\theta_R$, $R\times C$ and $\theta_R\times C$.

where $V(0)$ denotes the value of the project with flexibility, that is

$$V(0) = \sum_{j=0}^{N} \exp(-rj\Delta)\overline{V}(j\Delta).$$

4 Data and parameter estimates

The data used for the estimation of a taxi driver's cash-flows are obtained from the book *"Papers del Taxi. Reestructuració"* published by the IMET. This book summarizes a study of the sector in Barcelona during 1995 coinciding with the debate on how to restructure the sector. It provides estimates of costs and income of a representative taxi driver. Therefore, for convenience, we will make 1995, denoted as $t = 0$, the base year of our study. The data in that book will be complemented by estimations of the future evolution of variables such as income, fares or costs of inputs – mainly gas oil – from several sources that we will outline next.

Finally, at the end of this section we will present the estimates for the net surplus value from our model. We will call this first estimation of the surplus as the benchmark value. In the following section, we will do extensive sensitivity analysis to show the robustness of the benchmark.

4.1 Costs

4.1.1 Fixed costs and non-stochastic variable costs

We decompose the non-stochastic costs in three components: the costs related to the purchase of the car, $F_1(t)$, the rest of the fixed costs, $F_2(t)$, and the variable but deterministic costs, $F_3(t)$. We will later consider the costs of fuel.

According to the study undertaken by the IMET, the average price of a new taxicab in 1995, all taxes included, is €13,642.97.[16] The average duration is considered to be 7 years, and for simplicity we assume that depreciation is linear and constant over these 7 years, i.e. €1,949.99 per year. We will further assume that the car is paid during this 7 year period. Installments will amount to €1,949.99 a year together with the interest expenses which correspond to €409.29 a year. Hence, the total financial costs, that include amortization and interest payment amount in 1995, i.e. $F_1(0)$, to €2,359.28.

The fixed costs of running a taxi include, among other concepts, taxes, insurance, social security, value added taxes (VAT),[17] the income tax, parking and surveillance, and management and administration. Note that these kind of costs are independent of the distance driven by the taxicab and will be denoted as $F_2(t)$. All together they represent a cost $F_2(0)$ of €5,808.89. Finally, we compute the non-stochastic variable cost component in our model that collects the expenses on lubricants, tires and maintenance and repairs. Because they depend on the distance

[16] The study was performed in pesetas. For convenience, we have converted them into euros using the actual fixed exchange rate.

[17] Because the government cannot verify the revenue of the taxi driver, the VAT paid is for the most part established independently of production.

driven and not on the number of customers we can consider them as fixed.[18] Its estimation in the IMET study leads to an initial figure for 1995, $F_3(0)$, of €1,758.89 per year. The costs of the details corresponding to these three components for the base year are summarized in Table 1.

Table 1: Cost components in 1995 (in euros)

Components of $\mathbf{F}_1$	$\mathbf{F}_1(0)$	%
amortization	1,949.99	82.65
interest payment	409.29	17.35
Total	*2,359.28*	*100*
Components of $\mathbf{F}_2$	$\mathbf{F}_2(0)$	%
driving tax	164.81	2.84
vehicle insurance	1,341.75	23.10
social security	2,069.56	35.63
value added tax	360.61	6.21
income tax	841.42	14.48
parking and surveillance	865.46	14.90
management and administration	165.28	2.84
Total	*5,808.89*	*100*
Components of $\mathbf{F}_3$	$\mathbf{F}_3(0)$	%
lubricants	155.46	8.84
tires	415.13	23.60
maintenance and repairs	1,188.30	67.56
Total	*1,758.89*	*100*

Summing up, the total fixed costs and non-stochastic variable costs, denoted as $F(t)$, is the sum of the above three components. We assume that $F(t)$ grows at the inflation rate. We will set an inflation rate of 2.7% per year from now on,[19] then $\pi(t) = 0.027$ for $t \geq 0$. To simplify, it will be denoted as π this annual constant expected inflation rate.

4.1.2 Fuel costs

The cost of fuel $C(t)$ is the largest component of the variable costs. In 1995, it represents a cost $C(0)$ equal to €1,865.76. Next, we will estimate the parameters of the log-price equation (17)

[18] An important assumption in this case is that the distance is independent of variables such as the cost of fuel. It can be argued that if the cost of fuel increases taxi drivers spend less on the road. This feature could be incorpored in the model without difficulty. However, we lack data on the elasticity of this distance with respect to the price of gas oil. For this reason, we consider distance to keep constant.

[19] This is the target value under the EMU proposal in 1998.

which are necessary to implement equation (27). The parameters in (17) are under the risk-neutral measure. Notice that the risk premium ϕ_s is not observed through the spot diesel fuel price but it is implied in the futures price where the underlying asset is the gas oil commodity.

Following Schwartz (1997), the theoretical futures (or forward) price at time t where the underlying asset is driven by the one factor model defined in equation (17) and having maturity T is

$$F(t,T) = \exp\left[e^{-k_s T} X(t) + \alpha^* \left(1 - e^{-k_s T}\right) + \frac{\sigma_s^2}{4k_s}\left(1 - e^{-2k_s T}\right)\right]. \tag{30}$$

Let $\Phi \equiv (\alpha^*, k_s, \sigma_s)$ denote the vector of parameters to be estimated and consider $X(t)$ an unobservable state variable as in Schwartz (1997). Then, we estimate Φ by using the methodology given in Cortazar and Schwartz (2003) which is a procedure that is not based on the Kalman filtering method implemented in Schwartz (1997) but through a calibration procedure.[20] The estimation period goes from 12/24/90 to 12/25/95 for weekly observations with a sample size $m = 262$. The number of future contracts corresponding to every observation is always the same and equal to $n = 7$, so the total sample size is 1834. The gas oil futures contracts used are traded in the International Petroleum Exchange (IPE) of London and they are denoted as F1 to F7, where F1 is the contract closest to maturity, F2 is the second one closest to maturity and so on. The descriptive statistics for these contracts are shown in Table 2 for the above sample period.

Table 2: Gas oil futures contracts (in euros)

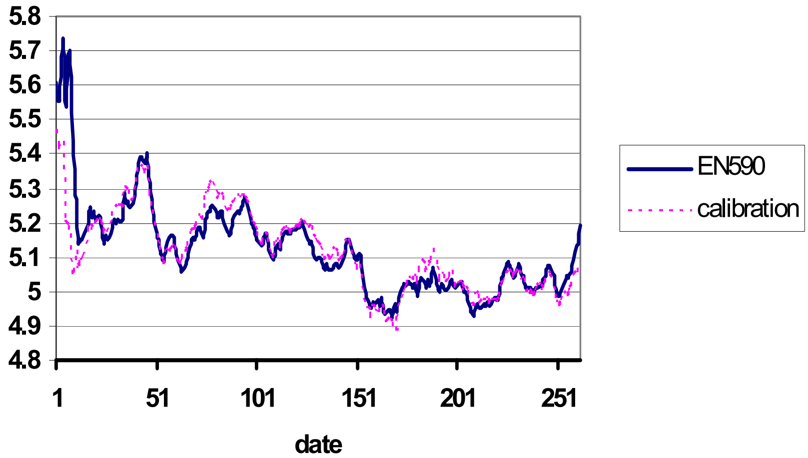
Futures	Mean price (s.e.)	Mean maturity in years (s.e.)
F1	167.68 (19.94)	0.038 (0.024)
F2	167.48 (18.71)	0.120 (0.024)
F3	167.54 (17.55)	0.202 (0.024)
F4	167.72 (16.46)	0.285 (0.024)
F5	167.69 (15.23)	0.367 (0.024)
F6	167.56 (14.11)	0.449 (0.024)
F7	167.38 (13.17)	0.531 (0.024)

The parameter estimates from equation (38) in Appendix B are $\alpha^* = 5.099$, $k_s = 0.782$ and $\sigma_s = 0.218$. For the simulation of equation (27) we need to obtain the risk premium ϕ_s. Since $\phi_s = \alpha - \alpha^*$, we need to obtain previously an estimate of α. We consider as a proxy for the underlying asset of gas oil futures the EN590 NWE FOB Barges gas oil spot from Platts.[21] We consider again both weekly observations and the same sample period as before. We estimate the exact discretization corresponding to (4) which follows an AR(1) representation. The result is α equal to 5.092 and therefore $\phi_s = -0.007$. In Figure 1 we can compare the evolution of the

[20] The estimation procedure to obtain the parameters can be seen in Appendix B.

[21] See for more details the web page: www.platts.com

Figure 1: Evolution of EN590 spot vs spot from calibration



EN590 spot price against the spot price series obtained through the calibration method (see equation (37) from Appendix B). We can see that the calibrated series follows very closely the original one, with a correlation between both series of around 83%. Note that the series from calibration through the futures model (30) does not need to replicate exactly the behavior of EN590 since the last one is not really the underlying asset of the futures but a proxy one.

4.2 Revenues

4.2.1 Initial value of revenues

The IMET estimated the average income of a taxi driver in 1995 using several methods. The result is that an appropriate estimation of a taxi driver's daily income is €93.16. Considering several samples, it is assessed that each taxicab makes on average about 28 trips a day. Since an average driver works around 221 days a year, the revenue in 1995, i.e. $R(0)$, is roughly €20,584.66.

4.2.2 Parameters of RGDP equation

Equation (25) requires estimates of the parameters that determine the real GDP growth rate in (10). We take the real GDP series from the Quarterly Spanish National Accounts[22] (QSNA) with 1986 constant prices. Our sample includes 104 observations corresponding to the period from 1970:1 to 1995:4. The real GDP return series $\theta_q(t)$ is created as $\theta_q(t) \equiv \ln RGDP(t) -$

[22] These data series and most of the ones used in the rest of the section are obtained from the TEMPUS database of the Spanish National Statistics Institute. See the webpage: www.ine.es

$\ln RGDP\,(t-1)$. The descriptive statistics for the $\theta_q\,(t)$ series can be seen in Table 3. The Jarque-Bera test does not reject the null hyphotesis of being a Normal distribution and we can appreciate that the time series is rather symmetric. The correlations, denoted as $\rho\,(i)$, tend to decrease quickly.

Table 3: Descriptive statistics of $\theta_q(t)$

Mean	0.007	$\rho\,(1)$	0.856
Median	0.007	$\rho\,(2)$	0.677
Std. Dev.	0.006	$\rho\,(3)$	0.563
Skewness	0.233	$\rho\,(4)$	0.477
Kurtosis	2.980	$\rho\,(5)$	0.350
Jarque-Bera (p-value)	0.627	$\rho\,(6)$	0.260

Since the exact discrete equation corresponding to the continuous equation defined in (10) is an AR(1) model with constant, we regress $\theta_q\,(t)$ on a constant and the one period lagged dependent variable with result shown in Table 4.

Table 4: LS with dependent variable $\theta_q(t)$

Independent variables	coefficient	std. error (s.e.)	p-value
c	0.001	5×10^{-4}	0.037
$\theta_q\,(t-1)$	0.858	0.051	0.000
R-squared	0.739	s.e. of regression	0.003

Given the estimates in the previous table, and considering $\Delta = \frac{1}{4}$, we can easily obtain the values of $\overline{\theta}$, k_θ and σ_θ through equation (25). The parameter values[23] are respectively $\overline{\theta} = 0.007$, $k_\theta = 0.613$ and $\sigma_\theta = 0.007$. Finally, the starting value for the last quarter of 1995, $\theta_q\,(0)$, is set to be 0.4% according to QSNA.

4.2.3 Revenue risk premium

We consider that the main source of risk in the revenue of the taxi driver is related to shifts in the local economy of the city. To the extent that the economy of Barcelona depends on the behavior of its financial service sector, we can approximate the risk of driving a taxi to the risk

[23] The parameters of RGDP equation are finally obtained by solving the following equations for $\Delta = 1/4$: $e^{-k_\theta\Delta} = 0.858$; $\overline{\theta}(1-e^{-k_\theta\Delta}) = 0.001$ and $\sigma_\theta\sqrt{\frac{1-e^{-2k_\theta\Delta}}{2\kappa_\theta}} = 0.003$.

of investing in any financial service firm. The average beta[24] of financial service firms is 0.817. The annual return for the Madrid Stock Exchange Index (IGBM) over the period 1994 to 1995 is 12.30%. The annual nominal interest rate r is obtained through 3-month interbank loans and it is equal to 9.4% in 1995. The market risk premium is 0.029 and so, the risk premium for the taxi revenues is $\phi_q = 0.817 \times 0.029 = 0.024$.

4.2.4 Volatility of revenues

The annual volatility of the variable $q(t)$ in (9) can not be estimated directly from the IMET study. So, it is approximated to the annual volatility for the series resulting as the difference between the growth series of the real gross value added (GVA) in market services from QSNA (base 1986) and the RGDP return series. Definitively, σ_q equals 0.6% given a sample holding the period 1970:2 to 1995:4.

4.2.5 Jump component in revenues

We would expect a perfect pass-through of changes in costs into changes in fares if taxi drivers had the power to choose prices. However, fares are the result of the negotiation by public authorities, in particular between the municipality and the taxi union. Since changes in fares do not occur on a regular basis, they will be captured through the Poisson process $N_p(t)$ with parameter λ_p. We assume that on average fares change once every five years, so the annual average number of jumps λ_p equals 0.2.

4.2.6 Parameters of deflator equation

The two parameters describing the drift and variance of the growth rate of the GDP deflator, β_D and σ_D, are easily obtained through equation (14). Given the quarterly GDPD series (base 1986) obtained from both the real and market price GDP, the result is $\sigma_D = 2.5\%$ for the sample standard deviation per year from the GDPD growth series corresponding to our period from 1970:2 to 1995:4. The sample mean GDPD annual growth rate is $\beta_D = 10.1\%$. Finally, we compute the starting value of the deflator growth rate $\varphi(0)$ to be 0.82% for the fourth quarter in 1995.

4.3 License price, alternative wage and other parameters

4.3.1 License price

We study the case of a taxi driver aged 35 who obtains a license in 1995 and so, he can exercise his job for 30 years. The license price in 1995, $L(0)$, amounts to €36,000 in 1995 and it increases every year according to the above mentioned expected inflation rate of π equals 2.7%. To simplify the exposition of the results, we assume in this example that the driver who

[24] This beta is an average of several betas. They have been estimated by considering the last 60 trading days in 1995.

buys the license is not financially constrained and so, he can pay the license without additional borrowing. The driver has the option at each point in time before retirement to sell the license at the fixed price mentioned above and exercise an alternative job for the remaining period until retirement.

4.3.2 Alternative wage

The IMET study estimates an implicit wage for the taxi driver at €11,419.23 a year in 1995. Assuming that the taxi driver is risk averse, we could estimate the certainty equivalent of his salary as an approximation to the alternative salary, denoted as $w(t)$. We use that as the minimum annual wage that he could accept in a different job. Given the revenue risk premium ϕ_q, the original wage corresponds to $w(0) = (1 - \phi_q) \times 11,419.23$ resulting in €11,145.17. We assume that growth rate for the alternative wage is the inflation rate π.

4.3.3 Correlation parameters

Finally, the remaining parameters that we need to determine their values are the correlations ρ_{sc}, $\rho_{c\theta}$ and $\rho_{\theta D}$. Since the sign of $\rho_{c\theta}$ is assumed to be negative as empirical evidence would suggest and also due to the negative strong impact of high oil prices in the economy measured through the behavior of the GDP growth rate, then $\rho_{c\theta}$ will take, for example, a value of -0.6. We will also give a high but positive value of ρ_{sc}. Take, for example, a value for ρ_{sc} equal 0.85. Finally, we will consider a value of $\rho_{\theta D}$ equal 0. Notice that all these correlations will be changed in the next section for the sensitivity analysis.

4.3.4 Simulation parameters

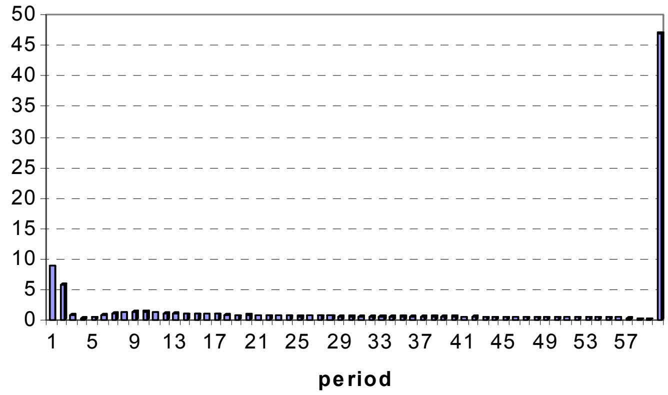
In all cases we use 50,000 simulations with half-yearly steps (periods), so that $\Delta = 1/2$, and up to a horizon of 30 years. As a result, $T = 30$ and $N = 60$.

4.4 Simulation results

Using both the data and parameter values described above, we obtain that the project value with flexibility, $V(0)$, is €417,690 and the taxi driver's net surplus, $\Gamma(0)$, is about €80,000.[25] In Figure 2, it is shown the evolution of the percentage of abandoned paths per period. We can appreciate two important spikes at the beginning, specifically during the first year (2 periods) in the future, i.e. year 1996. The total percentage of abandonments in 1996 amounts to 16.61%. The reason for this behavior could be explained in the low revenues that the taxi driver could earn because of his learning process in his new job as a taxi driver. Finally, we can appreciate another spike, that would be the largest one (46.82%) in the last period. It means that almost half of the taxi drivers decide to maintain this job until his retirement.

[25] To be more precise, the net surplus is exactly €79,989.

Figure 2: Percentage of abandoned paths per period ($T = 30$)



5 Sensitivity analysis

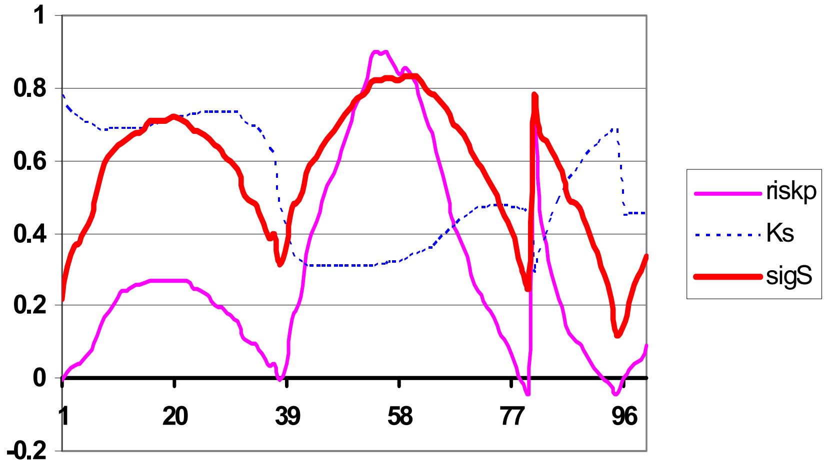
In this section we perform an exhaustive sensitivity analysis in order to measure the robustness of the net surplus value obtained in the previous section (also known as the benchmark value). The systematic analysis we carry out allow us, among other things, to compute the contribution of each of the variables we discuss on the surplus value.

5.1 Parameters of variable cost equation

We first study the sensibility of our results to changes in the parameters that govern the variable costs of the firm, and particularly, the price of the gas oil. These parameters, as shown in (27) are k_s, σ_s and ϕ_s.

In order to analyze the robustness of the estimates described in Section 4.1.2 for the whole sample from 12/24/90 to 12/25/95, we apply here the same methodology but using a different sample size each time under the following *rolling procedure*: in each estimation we reduce by one the number of observations, deleting the first one of the previous sample. That is, the first estimation corresponds to all the sample, while in the second we start from the observation of the date 12/31/90. The last estimation is done with data starting 12/26/94 so that we have at least one year of weekly data. As a result, we obtain a sample of 100 estimations where each has the estimates of the above three parameters for a certain sample period. We obtain the net surplus value when holding fixed the rest of the parameters in the model given its corresponding values in the last section. Figure 3 shows the evolution of the above three estimates through time where "riskp", "Ks" and "sigS" denote the parameters ϕ_s, k_s and σ_s respectively. A remarkable feature is the similarity in the behavior of ϕ_s and σ_s with a high correlation of about 0.85. The correlation between ϕ_s and k_s is about -0.67 and finally, the lowest correlation is between k_s and σ_s of -0.33. Table 5 shows the descriptive statistics of the

Figure 3: Evolution of variable cost parameter estimates



three series displayed in Figure 3 and also for the net surplus estimates. Note that the median of the net surplus series is €78,167 - higher than the mean but not too much - which is around the value obtained in the example from the last section, specifically €80,000. Meanwhile, the range corresponding to the net surplus values - going from €1,597 to €101,659 - shows a high dispersion.[26] Summing up, if we select the median as the candidate[27] of the estimation for the net surplus, then there is hardly difference with respect to the benchmark value of €80,000 from the last section.

Further analysis of the surplus series in Table 5 can be done by regressing the net surplus value (dependent variable) with respect to the above parameters (independent variables). The result of the regression, including the constant, has a R-squared of around 90% with all the parameters being significant.

The commovement of ϕ_s and σ_s mentioned before means that the information content of each parameter can be misleading. If we consider all cases of regressing any independent variable (ϕ_s, k_s, σ_s) against the rest of the independent ones including a constant, we find that when ϕ_s becomes the dependent one we obtain the highest R-squared, of around 89%. Denote ϕ_s^* as the error term of this last regression, that is the informational content of ϕ_s which is orthogonal to both k_s and σ_s. Table 6 displays a new regression with the surplus value as dependent variable but now with independent ones k_s, σ_s and ϕ_s^*. The results are as expected. The sign of ϕ_s^* is negative and since ρ_{sc} is positive under the benchmark framework (see Section 4.3.3), then

[26] Analysing better the surplus value series in Table 5, we find in the intervals $[0, 50)$, $[50, 100)$ and $[100, 102)$ the amounts of data - in thousands of euros - with percentages 36%, 55% and 9% respectively. So, most data can be found in the second interval that contains the median.

[27] The criterion is based on that the median is more robust than the mean.

Table 5: Descriptive statistics of ϕ_s, k_s, σ_s and net surplus

	ϕ_s	k_s	σ_s	$\Gamma(0)$ (in €)
Mean	0.294	0.525	0.571	60,237
Median	0.225	0.474	0.611	78,167
Maximum	0.897	0.782	0.835	101,659
Minimum	-0.045	0.294	0.118	1,597
Range	0.942	0.488	0.717	100,062
25th percentile	0.086	0.356	0.445	11,632
75th percentile	0.434	0.689	0.708	90,324
Std. Dev.	0.272	0.164	0.182	1,538

Table 6: LS with net surplus (in thousands of euros) as dep. var.

Indep. var.	coefficient	std. error. (s.e.)	p-value
c	-10.813	6.531	0.000
ϕ_s^*	-98.158	12.884	0.000
k_s	187.623	7.578	0.000
σ_s	-48.191	6.822	0.000
R-squared	0.907	s.e. of regression	11.650
Adj. R-squared	0.904	sample size	100

an increase in the risk-premium makes the surplus value to decrease due to the drift of the variable cost equation (18) is higher. Similarly, the negative sign of σ_s reflects the fact that higher volatility implies higher costs and so both lower cash-flows and lower surplus value.

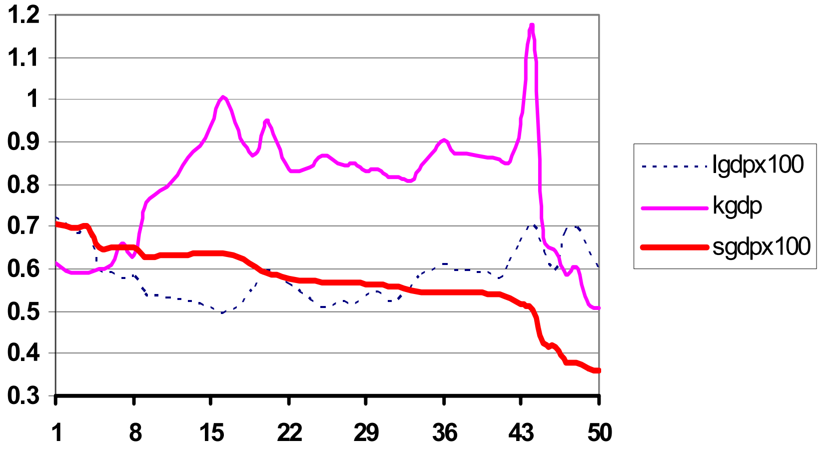
Finally, in Figure 3 we can appreciate that the volatility (sigS) shows a time-varying pattern and also, the commovement behavior of the volatility and the risk-premium (riskp). This could suggest a theoretical model for futures prices containing both stochastic volatility and a time-varying risk-premium as a function of the volatility. Summing up, computing futures prices under the well-known Heston's framework.[28] However, this extension is beyond the scope of this paper.

5.2 Parameters of RGDP equation

For the process governing the changes in the GDP we carry out the same study as before. Following a rolling procedure for the AR(1) structure established for RGDP in Section 4.2.2, we finally obtain a sample of fifty AR(1) estimates and transform them into the corresponding

[28] See Heston (1993).

Figure 4: Evolution of RGDP parameter estimates



parameter estimates of equation (25) to run the Monte Carlo simulation. We start in the first simulation with observations from 1970:2 to 1995:4, and going from 1982:3 to 1995:5 in the last one. Figure 4 shows the evolution of the parameters through time where "lgdp", "kgdp" and "sgdp" denote respectively the parameters $\overline{\theta}, k_\theta$ and σ_θ, meanwhile Table 7 presents the corresponding descriptive statistics.

Note that the volatility series (sgdp) shows a decreasing trend in Figure 4. The correlation values are -0.44, -0.23 and 0.13 corresponding respectively to $\left\{\overline{\theta}, k_\theta\right\}$,$\left\{\overline{\theta}, \sigma_\theta\right\}$ and $\left\{k_\theta, \sigma_\theta\right\}$.

Table 7: Descriptive Statistics of $\overline{\theta}$, k_θ, σ_θ and net surplus

	$\overline{\theta} \times 10^2$	k_θ	$\sigma_\theta \times 10^2$	$\Gamma(0)$ (in €)
Mean	0.581	0.796	0.570	70,059
Median	0.580	0.838	0.569	70,129
Maximum	0.720	1.173	0.705	81,736
Minimum	0.495	0.506	0.361	64,017
Range	0.225	0.667	0.344	17,719
25th percentile	0.526	0.662	0.543	65,957
75th percentile	0.597	0.871	0.633	71,689
Std. Dev.	0.062	0.139	0.085	4,643

Table 7 also presents the descriptive statistics of the net surplus series where each observation is obtained keeping all the parameters fixed except those from the RGDP equation (25) with different values - according to the above rolling procedure - per surplus value estimate. Note that the median value of €70,129 is lower than our benchmark value of €80,000. The series is rather symmetric. It shows a lower dispersion - see the standard deviation - than the corresponding one in Table 5. Finally, note that the maximum value of €81,736 is relatively close to the benchmark. We also study the effect on the surplus value of changes in the parameters. The corresponding regression – including also a constant term – shows a 99% R-squared coefficient. However, most of the explanatory power arises from $\overline{\theta}$, with a positive sign, while k_θ is not significant.

5.3 Correlation parameters

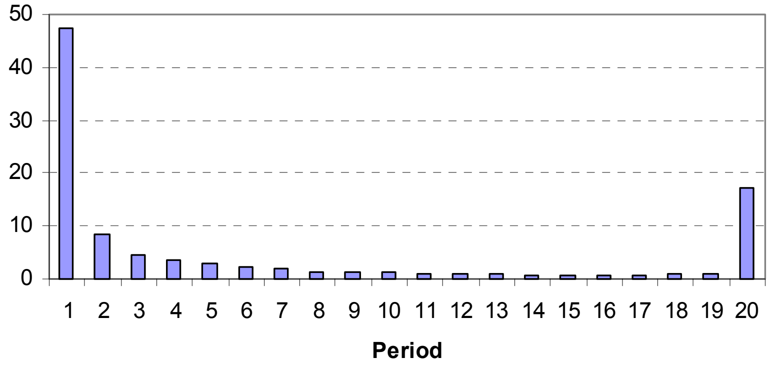
In this section we study how the net surplus changes when modifying in each surplus estimate only one correlation parameter value from the set $\{\rho_{sc}, \rho_{c\theta}, \rho_{\theta D}\}$ while the rest of the parameters are taken fixed from the benchmark case. The grid of values that we consider for any correlation goes from -1 to 1 with steps of size 0.02, so for each correlation we will have a surplus series of 101 observations. Table 8 shows the descriptive statistics corresponding to the net surplus series obtained in each case, that is one series per one correlation sensitivity analysis, which are denoted as $\Gamma_{\rho_{sc}}(0)$, $\Gamma_{\rho_{c\theta}}(0)$ and $\Gamma_{\rho_{\theta D}}(0)$ corresponding to ρ_{sc}, $\rho_{c\theta}$ and $\rho_{\theta D}$ respectively.

Table 8: Descriptive statistics for net surplus

	$\Gamma_{\rho_{c\theta}}(0)$ (in €)	$\Gamma_{\rho_{sc}}(0)$ (in €)	$\Gamma_{\rho_{\theta D}}(0)$ (in €)
Mean	77,880	82,246	80,159
Median (Q_2)	77,568	82,346	79,964
Maximum	82,405	84,792	85,730
Minimum	75,080	79,906	75,039
Range	7,325	4,886	10,691
25th percentile (Q_1)	76,062	80,895	77,315
75th percentile (Q_3)	79,531	83,404	82,950
Std. Dev.	2,132	1,473	3,131
Std. Dev./ Mean	0.027	0.018	0.039
$(Q_3 - Q_1)/Q_2$	0.045	0.030	0.070

Notice that the mean and median values corresponding to the three series are very similar to the ones under the benchmark case of €80,000. We can also appreciate that $\Gamma(0)$ is most sensitive with respect to $\rho_{\theta D}$ (see the relative dispersion measures located in the last two rows of Table 8), meanwhile ρ_{sc} does not seem to have a significant impact on the surplus value. Finally, regressing each surplus series against a constant and the correlation series we obtain that ρ_{sc} and $\rho_{c\theta}$ have a negative effect on the value of the surplus while the effect of $\rho_{\theta D}$ is positive. The corresponding R-squared measures are above 97% in all three regressions.

Figure 5: Percentage of abandoned paths per period ($T = 10$)



5.4 Time horizon

We finally measure the impact on the surplus value by considering only a decrease in the time horizon of driving a cab. Table 9 shows the effect of changing only the number of years from 30 to 5 under the benchmark situation. As expected, the surplus value decreases when the number of years is reduced and the probability of abandonment in the first period increases.

Table 9: Impact of time horizon on net surplus

T (years)	30	25	20	15	10	5
$\Gamma(0)$ (in €)	80,000	55,670	33,080	13,310	-70	-3,700
1st period (% aband.)	8.86	11.04	14.53	22.51	47.41	93.93
last period (% aband.)	46.81	46.81	45.31	38.16	17.29	1.34

In Figure 5 we present the evolution of the percentage of abandoned paths per period for a horizon fixed to 10 years. Compared with Figure 2 the main difference is that with a shorter horizon the agent abandons during the first periods. The reason is that for a short horizon, high initial costs or low initial revenues are enough to make the job nonprofitable.

Definitively, if we consider all the different sensitivity studies carried out in this section, we can conclude that the correlation analysis is the least sensitive showing net surplus values close to the benchmark value.

6 Extensions

In this section, we provide an extension of the framework introduced in Section 2. We extend the process for changes in quantities in (9) to allow for a more general case where the drift, that is a growth rate, is modelled now as a mean-reverting process to the growth rate of the real GDP represented as $\theta_q(t)$. If we denote, under this more general framework, the drift of the process for $dq(t)/q(t)$ as $d\beta(t)$ instead of $d\theta_q(t)$, we can assume

$$d\beta(t) = k_\beta\left(\theta_q(t) - \beta(t)\right)dt + \sigma_\beta dW_\beta(t)$$

$$d\theta_q(t) = k_\theta\left(\overline{\theta} - \theta_q(t)\right)dt + \sigma_\theta dW_\theta(t)$$

$$dW_\beta(t)\, dW_\theta(t) = \rho_{\beta\theta}$$

where $k_\beta, \sigma_\beta > 0$. The above model represents a bivariate Ornstein-Uhlenbeck process. The interpretation is that for fixed fares, the taxi driver's revenues does not depend exactly on the evolution of the RGDP growth rate. Instead, they are subject to a (idiosyncratic) shock that fades away over time. This new growth rate $\beta(t)$ should be positively correlated with $\theta_q(t)$, that is why we assume at first that $\rho_{\beta\theta} > 0$. Since we have no information a priori about how to estimate the parameters k_β and σ_β, we will assume that k_β equals k_θ and σ_β equals σ_θ. For values of $\rho_{\beta\theta}$ equal 0 and 0.8, we obtain net surplus values of €78,280 and €82,110 respectively. In general, a higher correlation between the growth rate of income and growth rate of GDP, which implies a smaller uninsurable idiosyncratic risk, increases the surplus value. In any case, the change in the surplus value is is around the benchmark value of €80,000.

7 Concluding Remarks

The taxi-sector is particularly prone to discussions on whether it should be reformed because it includes a variety of - sometimes contradictory - regulations. Many reforms have been focussed on liberalizing prices or making entry easier while at the same time trying to guarantee a certain income for the taxi drivers. Others, like our present study, have been focussed on insurance for the taxi drivers against fluctuations in the return from their job.

We show in this case how such a reform can increase the net surplus value for insiders respect to a liberalized market for licenses. The total effect we compute is substantial. The net surplus we estimate is around €80,000 and this value is robust to the sensitivity analysis carried out.

The study also points out one of the problems of this kind of reforms: its sustainability. Although this fixed license price scheme provided substantial protection during the years of crisis in the sector it also became inconvenient for the licenses once the sector recovered and the liberalized price of the license increased. The consequence of this difference was the pressure to abandon the mechanism.

Finally, a year after the liberalization of the sector, anecdotal evidence suggests that the price of the license in Barcelona has increased substantially and is barely similar to the prices in other Spanish cities.

References

[1] Bergstrom, A.R. (1984), "Continuous time stochastic models and issues of aggregation over time", in *Handbook of Econometrics,* vol. 2, 1145-1212, ed. by Z. Griliches and M.D. Intriligator, North Holland, Amsterdam.

[2] Cairns, R.D. and C. Liston-Heyes (1996), "Competition and regulation in the taxi industry," J*ournal of Public Economics*, 59, 1-15.

[3] Cortazar, G. and E.S. Schwartz (2003), "Implementing a stochastic model for oil futures prices", *Energy Economics* 25, 215-238.

[4] Dixit, A. and R. Pindyck (1994), *Investment under Uncertainty.* Princeton: Princeton University Press.

[5] D. Duffie (2001), *Dynamic Asset Pricing Theory.* Princeton: Princeton University Press, third edition.

[6] Heston, S. (1993), "A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options", *Review of Financial Studies*, 6, 327-343.

[7] Institut Metropolità del Taxi (1995), "Papers del Taxi. Reestructuració", EMT, Area Metropolitana de Barcelona.

[8] Longstaff, F.A. and E.S. Schwartz (2001), "Valuing American options by simulation: a simple least-square approach", *Review of Financial Studies,* vol. 14, no.1, 113-147.

[9] Schwartz, E.S. (1997), "The stochastic behavior of commodity prices: implications for valuation and hedging", *Journal of Finance* 52, 923-973.

[10] Schwartz E. S. (2001), "Patents and R&D as Real Options," working paper, UCLA.

[11] Schwartz, E.S: and M. Moon (2000), "Rational pricing of internet companies," *The Financial Review* 36, 7-26.

[12] Teal, R. F. and M. Berglund (1987), "The impacts of taxicab deregulation in the USA". *Journal of Transport Economics and Policy* 21, 37-56.

A Appendix[29]

Equation (4) implies that

$$\begin{aligned}\int_{t-\Delta}^{t} dX(u) &= \int_{t-\Delta}^{t} k_s(\alpha - X(u))\,du + \int_{t-\Delta}^{t} \sigma_s dW_s(u) \\ &= k_s\alpha\Delta - k_s\int_{t-\Delta}^{t} X(u)\,du + \eta_\gamma(t)\end{aligned} \tag{31}$$

where $\eta_\gamma(t) \sim N(0, \sigma_s^2\Delta)$. Also, note that $\int_{t-\Delta}^{t} X(u)\,du$ is a flow variable in contrast to the stock variable $S(t)$ and it will be denoted as $\xi(t)$. Then, equation for $\gamma_\Delta(t)$ in (5) may be expressed as

$$\gamma_\Delta(t) = k_s\alpha - k_s\Delta^{-1}\xi(t) + \Delta^{-1}\eta_\gamma(t) \tag{32}$$

It is shown in Theorem 8 from Bergstrom (1984) that $\xi(t)$ satisfies an ARMA (1,1) process, i.e.

$$\xi(t) = \alpha\Delta\left(1 - e^{-k_s\Delta}\right) + e^{-k_s\Delta}\xi(t-\Delta) + \eta_\xi(t) \tag{33}$$

where $\eta_\xi(t)$ is defined as

$$\begin{aligned}\eta_\xi(t) &= -\sigma_s k_s^{-1}\int_{t-\Delta}^{t}\left(e^{-(t-u)k_s} - 1\right)dW_s(u) \\ &\quad -\sigma_s k_s^{-1}\int_{t-2\Delta}^{t-\Delta}\left(e^{-k_s\Delta} - e^{-(t-\Delta-u)k_s}\right)dW_s(u)\end{aligned} \tag{34}$$

which follows a MA(1) process verifying[30]:

$$E\left[\eta_\xi^2(t)\right] = \frac{\sigma_s^2}{2k_s^3}\left[(1+k_s\Delta)\,e^{-2k_s\Delta} + k_s\Delta - 1\right]$$

$$E\left[\eta_\xi(t)\,\eta_\xi(t-\Delta)\right] = \frac{\sigma_s^2}{2k_s^3}\left[3\left(1 - e^{-2k_s\Delta}\right) - 2k_s\Delta e^{-k_s\Delta}\right]$$

We can appreciate that the disturbances $\eta_\gamma(t)$ and $\eta_\xi(t)$ defined in (31) and (34) are correlated and satisfy that

$$E\left(\eta_\gamma(t)\,\eta_\xi(t)\right) = \sigma_s^2 k_s^{-2}\left[e^{-k_s\Delta} + k_s\Delta - 1\right]$$

and

$$E\left(\eta_\gamma(t-\Delta)\,\eta_\xi(t)\right) = \sigma_s^2 k_s^{-2}\left[1 - (1+k_s\Delta)\,e^{-k_s\Delta}\right]$$

[29] In order to shorten this appendix the proof of some results here are available to the authors upon request.

[30] We apply equations (60) and (61) in Theorem 8 from Bergstrom (1984) for the general case of a time interval of length Δ in order to obtain the variance and the first-order covariance of $\eta_\xi(t)$.

Finally, considering equations (32) and (33) it can be shown that $\gamma_{\Delta}(t)$ follows an ARMA (1,1) process:

$$\gamma_{\Delta}(t) = e^{-k_s\Delta}\gamma_{\Delta}(t-\Delta) + \epsilon(t)$$

where $\epsilon(t)$ is a MA(1) process such that

$$E\ \left[\epsilon^2(t)\right] = \frac{\sigma_s^2}{2k_s\Delta^2}\left[3 - k_s\Delta - (3 + k_s\Delta)\, e^{-2k_s\Delta}\right]$$

and

$$E\ \left[\epsilon(t)\,\epsilon(t-\Delta)\right] = \frac{\sigma_s^2}{2k_s\Delta^2}\left[1 - e^{-k_s\Delta}\right]$$

B Appendix

In this appendix we explain in detail how to apply the methodology in Cortazar and Schwartz (2003) based on calibration to obtain the parameters in equation (30).

The log price corresponding to (30) for a certain date t_i and maturity T_j can be expressed as

$$\ln F(t_i, T_j) = \vartheta_0(T_j) + X(t_i)\vartheta_1(T_j) \tag{35}$$

such that

$$\begin{aligned}\vartheta_0(T_j) &\equiv \alpha^*\left(1 - e^{-k_s T_j}\right) + \frac{\sigma_s^2}{4k_s}\left(1 - e^{-2k_s T_j}\right)\\ \vartheta_1(T_j) &\equiv e^{-k_s T_j}\end{aligned}$$

Meanwhile, $\ln F_m(t_i, T_j)$ represents the log observed futures market price to the corresponding log theoretical one. Define the log pricing error e_{ij} corresponding to a certain date t_i and maturity T_j as the difference between both log prices which can be expressed given (35) as

$$\begin{aligned}e_{ij} &= \ln F(t_i, T_j) - \ln F_m(t_i, T_j)\\ &= \Lambda(t_i, T_j) - X(t_i)\vartheta_1(T_j)\end{aligned} \tag{36}$$

where $\Lambda(t_i, T_j) \equiv \ln F(t_i, T_j) - \vartheta_0(T_j)$. Note that for an initial value of Φ we have a simple linear regression model without a constant where the dependent variable is $\ln \widetilde{F}(t_i, T_j)$, $\vartheta(T_j)$ as the independent one and $X(t_i)$ as the parameter to estimate. So, given a cross-section sample for a certain date we could obtain the state variable by using the standard least squares (LS) regression. If we repeat this procedure given a time series with cross-section data, we finally obtain a time series for the state variable given the same initial value set of Φ. So, the LS estimate of $X(t_i)$, denoted as $\widehat{X}(t_i)$, for a given date t_i is obtained as

$$\widehat{X}(t_i) = \left(\sum_{j=1}^{n(t_i)} \vartheta_1^2(T_j)\right)^{-1} \sum_{j=1}^{n(t_i)} \Lambda(t_i, T_j)\vartheta_1(T_j) \tag{37}$$

where $n(t_i)$ represents the number of contracts in date t_i. Summing up, the calibration consists of minimizing the square errors defined in (36) for the whole sample. Note that e_{ij} becomes a function of both Φ and the state variable $\{X(t_i)\}_{i=1}^m$ where m is the time series sample size. Substituting $X(t_i)$ by its estimate in equation (37), then e_{ij} will only depend onto Φ. So, we finally minimize the following problem:

$$\min_{\{\Phi\}} \sum_{i=1}^{m} \sum_{j=1}^{n(t_i)} \left(\Lambda(t_i, T_j) - \widehat{X}(t_i)\vartheta_1(T_j)\right)^2 \tag{38}$$