FROM STRUCTURAL ESTIMATION TO QUASI EXPERIMENT

Three essays in empirical industrial organization

by

Gábor Koltay

Submitted in partial fulfillment to the requirements for the Degree of Doctor of Philosophy

at

Central European University

Budapest, Hungary

Supervisor: Prof. Gábor Kézdi

Co-Supervisor: Prof. Konrad Stahl

March 2012

CENTRAL EUROPEAN UNIVERSITY
DEPARTMENT OF ECONOMICS

The undersigned hereby certify that they have read and recommend to the Department of Economics for acceptance the thesis entitled "From structural estimation to quasi experiment: Three essays in empirical industrial organization" by Gábor Koltay.

Dated: February 2012

I certify that I have read this dissertation and in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Chair of the Thesis Committee: ______________________
László Csaba

I certify that I have read this dissertation and in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Advisor: ______________________
Gábor Kézdi

I certify that I have read this dissertation and in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Co-advisor: ______________________
Konrad Stahl

I certify that I have read this dissertation and in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Internal Examiner ______________________
Miklós Koren

I certify that I have read this dissertation and in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

External Examiner ______________________
Philipp Schmidt-Dengler

I certify that I have read this dissertation and in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Internal Member ______________________
Andrzej Baniak

I certify that I have read this dissertation and in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

External Member

Ádám Szentpéteri

I certify that I have read this dissertation and in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

External Member

Gábor Békés

CENTRAL EUROPEAN UNIVERSITY
DEPARTMENT OF ECONOMICS

Author: Gábor Koltay
Title: From structural estimation to quasi experiment: Three essays in empirical industrial organization
Department: Department of Economics
Degree: PhD
Dated: March 2012

Hereby I testify that this thesis contains no material accepted for any other degree in any other institution and that it contains no material previously written and/or published by another person except where appropriate acknowledgement is made.

Signature of the author:

Gábor Koltay

ABSTRACT

The thesis consists of three studies in empirical industrial organization. The first chapter examines how consumers choose eco-labeled products. It argues that eco-labels transform ordinary products into impure public goods: next to the usual product characteristics they offer consumers the possibility to contribute to reduced environmental pollution. As such, eco-labeled products are policy experiments in the private provision of public goods. Accordingly, when estimating the effect of an eco-label the consumer choice model has to incorporate various explanations for the private provision of public goods: pure altruism, warm glow and conditional co-operation. It is shown that conditional cooperation implies a demand system with interdependent preferences that transforms the usual discrete choice model into a discrete game. This modified demand system is identified based on the assumption of no conspicuous consumption. The model is estimated for the German eco-label the "Blauer Engel" using a household panel sample of toilet paper purchases. The results show that conditional cooperation is important in explaining consumer choices, although the effect of the Blauer Engel label is close to zero on average.

The second chapter contributes to the disaggregated evidence about asymmetric price transmission. It studies how station-level retail prices respond to wholesale price changes in the Hungarian gasoline market. The estimates show that although

retail price changes are almost symmetric on average, there is a subset of stations that follow an asymmetric pricing strategy. Having a closer look at station characteristics reveals that asymmetric pricing is a brand property and that these brands have small market share (below 10%) and are not vertically integrated. Other observables, like the number or the types of competitors do not explain the asymmetric retail price response. These results imply that in the same local market there are firms that price symmetrically and firms that price asymmetrically. This finding does not support collusion and search based explanations of asymmetric price transmission, because these are based on market level interactions among firms and consumers. Instead, it points towards the role of adjustment costs as an explanation for asymmetric retail price responses. Moreover, the result that the number and the types of competitors does not explain asymmetric pricing lends additional support to the claim that pricing asymmetry does not necessarily imply collusive behavior.

The third chapter applies difference-in-differences methods to identify the price effects of simultaneous mergers in the Hungarian retail gasoline market and to separate the different effects on the prices of the buyer and seller firms and on the prices of their respective competitors. The difference-in-differences approach exploits variation in the presence of merging firms across local markets to form different treatment-control group pairs in order to estimate separate effects for each type of firms affected by the mergers. This ex-post evaluation shows that both mergers resulted in a significant and positive but economically negligible price effect. Separating the simultaneous merger effects also reveals that the first merger affected only the prices of buyer firm's stations, the second had an effect on the prices of seller's stations and of its competitors. These results are not sensitive to the assumed dates when the mergers effectively change the firms' pricing policy.

Acknowledgements

First and foremost, I would like to express my gratitude to my supervisor, Gábor Kézdi, who patiently guided me from vague ideas to a finalized dissertation. Similarly, I am deeply indebted to Konrad Stahl. The first chapter of the thesis would still be a research plan without his belief in the project and he also provided invaluable help in accessing the data.

I owe special thanks to Gergely Csorba, who, besides coauthoring the third chapter, was always ready to discuss the problems I encountered during the writing of the thesis. Likewise, I would like to thank my other coauthor Dávid Farkas for his invaluable help throughout our joint research. I am also thankful for a number of former colleagues at the Institute of Economics and fellow PhD students, especially Zsombor Cseres-Gergely, Anikó Bíró and Péter Harasztosi for helpful discussions and companionship. I am indebted to John Earle and Álmos Telegdy, from whom I learned the nitty-gritty of empirical work, while being their research assistant at the CEU Labor Project. I would also like to thank Márton Csillag who always managed to summarize complicated arguments in a short sentence, and Matteo Salto for his help during the finalization of the thesis. I owe thanks to Katalin Szimler for all her patience and work as well. I am thankful to participants and discussants at various conferences and workshops for their comments, especially to Randy Filer.

A significant part of the thesis research was carried out while I was a junior researcher at the Institute of Economics, of the Hungarian Academy of Sciences.

The gasoline retail market research project was funded by the OTKA grant 73777. I owe gratitude to RAL Deutsches Institut für Gütesicherung und Kennzeichnung and GfK Germany for generously providing the data for the first chapter, and the GDN Eighth Annual Regional Research Competition for funding the research.

Last but not least, I would like to thank my family for their patience. I am especially grateful to Anna Horváth for reading through all the chapters and whose support always helped me through the critical moments.

Contents

Introduction

The thesis consists of three empirical studies in the field of empirical industrial organization and uses econometric tools to analyze the behavior of consumers and firms. The first chapter looks at the demand for eco-labeled products, with the aim of describing consumer behavior in these situations and estimating the willingness to pay for the German eco-label 'Blauer Engel'. In contrast to this demand side study, the following two chapters focus on the supply side and analyze firms' pricing decisions. The second chapter compares how gasoline stations change prices in increasing and decreasing wholesale price periods and tests the popular belief that retail prices increase faster and decrease slower in the respective periods. The analysis focuses on how such asymmetric pricing behavior varies with station characteristics with a view to differentiate among various existing explanations. Finally, the third chapter analyzes two, nearly simultaneous mergers and their effect on prices. The objective here is to differentiate among the different type of price effects predicted by economic theory.

The variety of topics in the thesis implies that the econometric methods used in the thesis cover a broad range as well. The first chapter is a study in structural estimation: a formal model of consumer behavior is derived and the parameters of this behavioral model are estimated. In the empirical industrial organization literature this structural approach is the most wide-spread, given its strong links to theory. The econometric method of the second chapter is most closely related to classical reduced form analysis: a flexible specification is estimated for a simple relationship of retail and wholesale prices and the results are compared to predictions from economic theory. The chapter goes one step further, however, and also discusses how far the estimated relationship can have a structural interpretation. Finally, the econometric approach of the third chapter follows the program evaluation literature and is based on the quasi experimental nature of mergers.

A major contribution of the thesis lays in the unique micro-datasets that were collected and assembled for the analysis. Chapter one is based on a German consumer-level panel data provided by the market research firm GfK Germany. This dataset is merged with information about the German eco-label 'Blauer Engel' made available by RAL, the German eco-labeling agency. Combining these two datasets gives unique information about the consumption of Blauer Engel-labeled products. The second and the third chapters are based on a station-level panel about the Hungarian gasoline retail market. This dataset was collected by the author from the online price comparison website “holtankoljak.hu”.

The contribution of the first chapter is threefold. First, it shows that eco-labels imply a specific demand structure where consumer choices are interdependent. Building on the model of Kotchen (2006) I argue that eco-labeled products are impure public goods, because the eco-label is de facto a tool that makes contribution to a public good (reduced environmental pollution) possible. I extend this model by incorporating the result from experimental economics (summarized in Fehr and Fischbacher, 2003) that public good contributions are usually explained by norms and reciprocity as well as by egoistic motivations (warm glow giving) or pure altruism. I show that incorporating such behavioral motives into the consumer choice model results in a demand system where the individual is influenced by the average expected choices of the others.

The second contribution of the chapter is that it suggests ways to identify the effect of other consumers on individual choices. As it is shown in the literature on discrete choice models with strategic interactions (Manski, 1993 and Brock and Durlauf 2001, 2003) as well as by the literature on static games (Bajari, Hong, Krainer, and Nekipelov, 2010, Seim, 2006 and Sweeting, 2009), the demand system with interdependent preference suffers from an endogeneity problem. I show that the standard restrictions used in discrete choice demand models are sufficient to solve this endogeneity. In addition, I also use geographic heterogeneity to separate the endogenous demand effect from the eco-labeling decisions made by firms.

The third contribution of the chapter is that it gives an estimate of the willingness to pay for the Blauer Engel eco-label, based on the unique consumption data. The estimates show that expectations about fellow consumers' behavior matter: in low expectation environments consumers are willing to pay 30% less, while in high purchase environments they are willing to pay 30% more for eco-labeled products. These results suggest that the success of such "self-regulation" is heavily dependent on the norms and beliefs of market participants.

The contribution of the second chapter is that it provides evidence on how asymmetric pricing behavior depends on observable characteristics of gasoline stations. Moreover, it shows how these findings can be used to differentiate between existing explanations for asymmetric pricing. This is made possible by the station-level micro dataset and there are still only a few studies that provide such evidence (for example Verlinda, 2008 and Faber, 2009). The chapter demonstrates that average pricing behavior might mask important heterogeneities and that pricing decisions should be analyzed at firm level.

The estimates show that although retail price changes are almost symmetric on aggregate, there is a subset of stations that follow an asymmetric pricing strategy. Station characteristics reveal that asymmetric pricing is a brand property: there are four brands that change prices with sizeable asymmetry. The brands that price asymmetrically have small market share (below 10%) and are not vertically integrated. These results imply that in the same local market some firms price symmetrically and some price asymmetrically.

This finding differentiates among theoretical models of asymmetric pricing. It does not support collusion and search based explanations that depend on market level interactions among firms and consumers. Instead it points towards the role of adjustment costs as an explanation for asymmetric retail price responses. Moreover, the result that the number of competitors and type of competitors does not explain asymmetric retail price responses lends additional support to the claim that pricing asymmetry does not necessarily imply collusive behavior. From a competition policy perspective it is important to determine whether the asymmetric retail price response is a result of anti-competitive behavior of firms or it can be explained by firms' costs and consumer behavior. The results support the latter.

The third chapter is joint work with Gergely Csorba and Dávid Farkas. It shows how the difference-in-differences method can be used to separate heterogeneous merger effects. The first contribution is that different effects are estimated for the two simultaneous mergers, for buyer and seller firms in the respective mergers and also for the competitors of the merging firms. Separating these merger effects enables one to test some important predictions of academic and antitrust literature, which argue that a merger can result in different price changes for different firms, depending on their role in the merger. First, the most robust prediction is that a merger will result in a larger change in merging firms' pricing than in competitor firms' pricing as the former can fully internalize the effect of eliminating the competitive constraint (externality) the two firms had on each other before the merger. Second, in mergers with local markets, a larger price increase is expected on markets where both merging firms are present (or are closer competitors to each other), since the merger removes a direct competitive constraint between their respective outlets. Third, a merger might have a different effect on the two firms involved, as the business policies and supply conditions of the firms will likely converge towards each other, and the change is usually conjectured to be larger for the case of the acquired firm than for the buyer firm.

The base result is that neither merger contributed substantially to retail price increases, as all estimated price changes are less than one percent. Most importantly, the two mergers had different effects on the merging firms depending on their role in the merger. These differences are broadly in line with the theoretical predictions. For the Agip/Esso merger, there are significant effects on the pricing of the acquired Esso stations and their competitors, and the price change is larger at Esso stations than at competitors' stations (although the difference is not significant). For the Lukoil/Jet merger own effects are larger than competitor effects, but a significant effect is found only for the buying firm's stations.

The second contribution of the chapter is that it discusses in detail the validity of difference-in-differences methods for merger effect estimation. An increasing number of studies (for example Hastings, 2004) began to use this estimation method, but they usually do not discuss whether the assumptions of this method are satisfied or not. However, this question is important, because mergers are not standard quasi experiments for the following reasons. First, most likely the merger decisions are endogenous since they are not independent of retail price decisions. Second, in most mergers only the clearance date is known, but not the date of the actual behavioral change. Therefore the definition of the periods before and after the merger depends on the judgment of the researcher and results can be sensitive to this judgment if the time series is not long enough. Third, through the price equilibrium the merger affects also the competitors of the merging parties. These market interaction effects become problematic if there are more mergers taking place parallel in the same market: without further assumptions the difference-in-differences estimator is not valid. The chapter addresses each of these problems and shows the sensitivity of the results to them.

CHAPTER 1

How much does your environment matter? Estimating the effect of the Blauer Engel eco-label

1.1. Introduction

Can voluntary actions of market participants successfully reduce environmental pollution? Eco-labels are based on the assumption that they can. These labels – the EU-wide Eco-label, the Blauer Engel in Germany or the Nordic Swan in the Scandinavian countries – indicate that the company adopted a voluntary standard in order to reduce the environmental impact of its product. This information enables consumers to pay directly for a cleaner environment by purchasing these eco-labeled products. In turn, producers are expected to label more and more products if they see that, everything else equal, consumers pay more for eco-labeled than for non-eco-labeled products.

It is uncertain, however, whether such "self-regulation" can work in practice. When the consumer purchases an eco-labeled product he does not only pay for the usual characteristics but also contributes to a public good: reduced environmental pollution. Therefore, the question is whether the private provision of such an impure public good is likely or not. Classical consumer theory suggests that there should be no positive contributions, because individual purchases are negligible compared to the size of the market and therefore individual actions have zero marginal environmental impact. Behavioral theory and economic models of moral behavior, on the other hand, suggest that positive utility is attached to the purchase of eco-labeled goods regardless of their environmental impact. This can be explained either by the strong personal motivation of the consumer (due to 'warm glow' giving or moral commitment) or by conditional cooperation that assumes consumers reciprocating the positive contributions by other consumers.

In markets with eco-labeled products one observes the choices of consumers, which allows the researcher to infer contributions to this public good. This paper provides such marketplace-based evidence by studying the effect of the German eco-label Blauer Engel using consumer choice data.

In order to estimate the effect of an eco-label on consumer decisions the alternative theoretical explanations for consumer choices over impure public goods has to be incorporated into an empirical choice model. The chapter shows that conditional cooperation implies that individual decisions will be interdependent: the effect of the eco-label will depend on whether other consumers are purchasing such goods or not; that is, whether they are contributing to the public good or not. Such a choice structure presents a challenge for identification since it will be difficult to separate individual decisions from average outcomes. The paper discusses the assumptions necessary for identification and applies the two step estimator of Bajari, Hong, Krainer and Nekipelov (2010) to overcome this problem.

The paper focus on the willingness to pay (WTP) for the eco-label when discussing the results. The WTP quantifies the value of the public good contribution in a straightforward manner: it gives the money value of the eco-label. As such, it is an important information for firms that plan to adopt eco-labels. If the WTP is negative or zero the eco-label adoption is not profitable and one cannot expect that eco-labeled products to spread in that market. Therefore, the WTP can also be a useful guide to the potential success of such voluntary regulation policies.

Although there are many experimental and contingent valuation studies on contributions to public goods, marketplace evidence is still quite rare. Contingent valuation studies, including Roe et al. (2001) on "green" electricity and Blend and Ravenswaay (1999) on eco-labeled food, usually provide evidence for positive willingness to pay for the environmental characteristic of the products. These studies also point out, however, that survey evidence is likely to overstate actual willingness to pay. Accordingly, marketplace evidence is more controversial. Bjørner, Hansen and Russell (2004) use Danish data and estimate that labels raise the willingness to pay by 13-18% percent of the retail price. On the contrary, Nimon and Beghin (1999) use apparel catalogue data and do not find any evidence for positive willingness to pay.

The results show that conditional cooperation is important: in low expectation environments consumers require a discount of around 30% in terms of the average price, while in high expectation environments they are wiling to pay roughly 30% for eco-labeled products. The overall average effect of the eco-label in this market is, however, close to zero and consumers who do not expect anybody else to buy eco-labeled product require a large discount. Both of these findings suggest that purely egoistic explanations for eco-labeled product choice is unlikely to be sufficient.

The paper is structured as follows. First, the consumer choice model is discussed showing how eco-labeled products can be interpreted as non-pure public goods and derives a discrete choice demand model based on it. Second, the identification problem is discussed. Third, the unique household panel dataset is introduced. Finally, the results are presented.

1.2. Consumer choice and eco-labels

1.2.1. Eco-labeled products as impure public goods

Consumption of eco-labeled products can be described by a discrete choice model. In such a model consumers choose among several products and purchase exactly one product at a time. Following Lancaster (1971) and McFadden (1974, 1981), discrete choice models describe products as bundles of characteristics, and consumer preferences are defined over these characteristics. Eco-labels can be incorporated in this framework by adding an additional characteristic to the product: its environmental impact that expresses the pollution emitted during the production, consumption and disposal of the product.

The environmental impact, however, is not a standard characteristic like color, packaging or size. First of all, it is not directly observable for the consumer, because much of the pollution is generated during the production and the disposal process. Eco-labels address exactly this imperfect information problem by signaling the otherwise unobservable environmental impact to consumers. Although the eco-label is awarded by fulfilling a well-defined list of criteria, consumers are unlikely to keep the exact details in mind. Therefore, the eco-label will be a noisy signal, but will indicate the firm's effort to reduce pollution.

The second specific aspect of the environmental impact of a product is that it is a public good (bad). The consumer usually lives far away from the factories that generate the pollution and makes her choice after the product was already manufactured. Therefore, consumers are only indirectly affected by the environmental impact of their own choices. It is, nevertheless, plausible to argue that if everybody would choose products with reduced environmental impact each consumer would be better off. These two aspects make the environmental impact of the product a classical case for a public good (bad). Given that products also have other characteristics than their environmental impact, eco-labeled products can be best described as impure public goods. This perspective also allows to regard "green markets" (markets with eco-labeled products) as policy experiments in the private provision of public goods.

Kotchen (2006) presents a model where choices over eco-labeled products are interpreted as choices over impure public goods. In his model the public good aspect is the reduced environmental impact signalled by the eco-label. I re-interpret the model of Kotchen (2006) in a discrete choice framework. Instead of looking at decisions about the quantities of a pure public good, a pure private good and an impure public good, I define consumer choices over a set of products that have different characteristics (including their environmental impact). Moreover, I also move beyond Kotchen's model by including other motivations than pure altruism in explaining consumer decisions.

A long line of research in behavioral economics (summarized for example in Fehr and Fischbacher, 2003) analyzes how subjects behave in public good experiments and provides strong evidence that people are not exclusively motivated by the immediate individual payoffs. Fehr, Fischbacher and Gächter (2001) show that experimental subjects are "conditional cooperators" who are willing to contribute to a public good if they expect others to do so as well. Fehr and Gächter (2002) and Rege (2004) point out the importance of punishments and the expression of social disapproval in maintaining such behavior. Theories that explain such behavior cover a wide range from social norms (Bernheim,1994, Brock and Durlauf, 2001 or Brekke, Kverndokk and Nyborg, 2003) to reciprocity (Sugden, 1984 or Fehr and Gächter, 2000).

"Warm glow giving" (Andreoni, 1990) or "unconditional cooperation" is another alternative explanation for private contributions to a public good. This explanation relies on the egoistic motivation of consumers and assumes that the act of giving itself is valuable to them, even if nobody else contributes to the public good. In order to capture these different motivations, I incorporate into the consumer choice model both conditional and unconditional cooperation as well as pure altruism.

1.2.2. The consumer choice model

Assume that there are $i = 1...N$ individuals who choose one product out of $j = 1...J$ alternatives. These choices are indicated by the binary variable y_{ij} that takes the value of one if individual i chooses product j and zero otherwise. A product type j is characterized by a vector of K characteristics, $\mathbf{x}_j = [x_{1j}, x_{2j}, ..., x_{Kj}]$, and a public good contribution g_j, which is implied by the consumption of product j. Accordingly, the public good contribution of individual i can be expressed as $g_i = \sum\limits_j y_{ij} g_j$.To keep notation simple, it is assumed that $\mathbf{x}_j$ includes income of i minus price of product j as well. The public good aspect of the alternatives implies that consumer choices are influenced not only by the product characteristics $\mathbf{x}_j$ and g_j but also by the total provision of the public good by other consumers, which will be denoted as $G_{-i} = \sum\limits_{n \neq i} g_n$. Besides these deterministic components, discrete choice models of consumer behavior account also for factors that are unobservable for the researcher but influence consumer choices by including a random utility component ε_{ij} with joint probability distribution $F(\varepsilon_{i1}, ..., \varepsilon_{iJ})$. Assuming that the random component enters additively, individual i's indirect utility function can be expressed as:

$$U_{ij} = V_i(\mathbf{x}_j, g_j, G_{-i}) + \varepsilon_{ij}. \tag{1.1}$$

In (1.1) the individual indirect utility, $V_i(\mathbf{x}_j, g_j, G_{-i})$, depends on other consumers' choices (y_{nj}) through $G_{-i} = \sum_{n \neq i} \sum_j y_{nj} g_j$. These choices are stochastic and the individual indirect utility, $V_i(\mathbf{x}_j, g_j, G_{-i})$ is probabilistic as well. Therefore the consumer's decision will be based on the expected representative utility:

$$V_i(\mathbf{x}_j, g_j) = \sum_{j_1}, ... \sum_{j_{i-1}}, \sum_{j_{i+1}}, ..., \sum_{j_N} \prod_{n \neq i} \Pr\left(y_{nj_n} = 1 | \mathbf{x}_j, g_j\right) V_i(\mathbf{x}_j, g_j, G_{-i}). \tag{1.2}$$

where j_k is the product choice index of individual k and $\Pr\left(y_{kj} = 1 | \mathbf{x}_j, g_j\right)$ is the probability that individual k chooses alternative j given the product characteristics. Then the consumer's utility maximizing choices will be described by the following probabilities:

$$P_{ij} = \Pr(\varepsilon_{ik} - \varepsilon_{ij} < V_i(\mathbf{x}_k, g_k) - V_i(\mathbf{x}_j, g_j), \forall k \neq j). \tag{1.3}$$

These probabilities form a demand system for the J different alternatives in which only utility differences among alternatives matter for the consumer's decision.

In order to derive such a demand system, one has to specify $V_i(\mathbf{x}_j, g_j, G_{-i})$ and therefore one has to consider carefully the explanations of consumer choices regarding the public good provision. Three cases will be considered:

(1) *Pure altruism*, which is the classical assumption in the public good literature. It implies that it is the aggregate contribution $G_j = g_j + G_{-i}$ that matters for the individual[1] resulting in the indirect utility specification: $V_i(\mathbf{x}_j, g_j + G_{-i})$. In this case individual i's contribution and the other consumers' contributions to the public good are perfect substitutes.

[1]Note that usually total contribution $Y = y_i + y_{-i}$, and does not vary with j. This is so because mostly it is assumed that the consumer optimizes over the quantity of her contribution y_i, implicitly through Y. In the current case, however the consumer can choose among j different type of contributions y_j by choosing one of the products.

(2) *Warm glow or strong personal ideals*, which imply that only the individual's own contribution g_j matters in her decision, yielding the indirect utility form: $V_i(\mathbf{x}_j, g_j)$. Warm glow giving is used to describe the case when decision-makers derive utility from the simple fact of giving (Andreoni, 1990) and do not care about the actions of other decision-makers.

(3) *Conditional cooperation*, which implies that individual contributions are complements to other consumers' contributions. There are several theories that belong to this category, for example conformity to social norms or reciprocity. These two can be captured with the utility formulation $V_i(\mathbf{x}_j, g_j \frac{G_{-i}}{N-1})$ that assumes that individual contributions and the average contribution are complements.

Formally, the impure altruism concept of Andreoni (1989) that specifies the individual utility as $V_i(\mathbf{x}_j, g_j + G_{-i}, g_j)$ is able to encompass all three indirect utility specifications. Therefore, the empirical model of this paper can be regarded as an application of this concept.

In the discrete choice demand literature a linear specification is assumed for $V_i(\mathbf{x}_j, g_j, G_{-i})$ in most cases (see for example Ackerberg et al, 2007). Although this formulation seems to be innocent, it implies relatively strong assumptions about the separability of the different behavioral motivations. In the pure altruism case the linear specification implies that the representative utility is:

$$V_i(\mathbf{x}_j, g_j + G_{-i}) = \beta_i \mathbf{x}_j + \theta_i (g_j + G_{-i}), \tag{1.4}$$

where vector $\boldsymbol{\beta}_i$ and scalar θ_i are individual specific parameters. Because only utility differences matter for the consumer's decision and $\theta_i G_{-i}$ is constant across alternatives the pure altruism case is observationally equivalent to the warm glow specification ($V_i(\mathbf{x}_j, g_j) = \boldsymbol{\beta}_i\mathbf{x}_j + \theta_i g_j$). To differentiate between the two cases complementarity of the product characteristics and the public good contributions has to be assumed. An example for such a specification for pure altruism is:

$$V_i(\mathbf{x}_j, g_j + G_{-i}) = \boldsymbol{\beta}_i\theta_i(g_j + G_{-i})\mathbf{x}_j \tag{1.5}$$

where the term $\boldsymbol{\beta}_i\theta_i G_{-i}\mathbf{x}_j$ will distinguish pure altruism from the warm glow specification. However, to keep the empirical specification comparable to previous studies, complementarity will not be introduced.

The third case, conditional cooperation, is the following under the linear utility assumption:

$$V_i(\mathbf{x}_j, g_j, G_{-i}) = \boldsymbol{\beta}_i\mathbf{x}_j + \theta_i G_{-i} g_j, \tag{1.6}$$

which can be distinguished from the other two cases without any further assumptions.

In order to incorporate all three cases of behavioral motivation, the following specification is suggested:

$$V_i(\mathbf{x}_j, g_j, G_{-i}) = \boldsymbol{\beta}_i\mathbf{x}_j + \theta_{1i} g_j + \theta_{2i}\frac{G_{-i}}{N-1} g_j. \tag{1.7}$$

The choice of average contribution by other consumers ($\frac{G_{-i}}{N-1}$) over total contributions reflects the assumption that it is reciprocity or social norms that steer individual behavior rather than preferences over the aggregate level of the public good. This is plausible since the reduction in environmental impact is difficult to quantify for the consumer.

The discussion so far assumed that the contribution of each product to the public good is directly observable, that is, that consumers have preferences directly over g_j and G_{-i}. As it was pointed out, however, consumers observe only a simple signal: whether product j is eco-labeled or not. This binary signal will be denoted by the indicator variable L_j that takes the value of 1 if j has the eco-label. Throughout the paper I will assume that the eco-label is a standardized environmental improvement across all products. Moreover, I will also assume that the signal it conveys to the consumer is fully credible. These assumptions imply that

$$g_j = \delta L_{j,}. \tag{1.8}$$

where δ is the standardized environmental improvement. Substituting this expression into (1.7) yields the following utility function:

$$V_i(\mathbf{x}_j, L_j, L_{-i}) = \beta_i \mathbf{x}_j + \theta_{1i}\delta L_j + \theta_{2i}\frac{\delta L_{-i}}{N-1}\delta L_j, \tag{1.9}$$

where $L_{-i} = \sum\limits_{n \neq i}\sum\limits_{j} y_{nj}L_j$ is the number of environmentally labelled products purchased by consumers other than i and $\frac{L_{-i}}{N-1}$ is the fraction of eco-labeled products purchased by consumers other than i. Taking expectations over other agents' choices, the expected representative utility for individual i is:

$$V_i(\mathbf{x}_j, L_j) = \boldsymbol{\beta}_i \mathbf{x}_j + \theta_{1i}\delta L_j + \theta_{2i}\delta^2 E\left(\frac{L_{-i}}{N-1}\right) L_j \tag{1.10}$$

where

$$E\left(\frac{L_{-i}}{N-1}\right) = \frac{1}{N-1}\sum_{k \neq i}\sum_{j} E(y_{kj})L_j = \sum_{j} P(y_{kj} = 1|\mathbf{x}_j, L_j, \boldsymbol{\theta})L_j \equiv P^L \tag{1.11}$$

is the aggregate probability that an eco-labeled product will be chosen given the product characteristics and the taste parameters $\boldsymbol{\theta} = [\boldsymbol{\beta}_i, \theta_{1i}, \theta_{2i}, \delta]$, which is nothing else than the expected market share of eco-labeled products. This probability will be denoted as P^L in the following. Because δ cannot be identified, it will be assumed to be 1, that is the θ_{1i}-s and θ_{2i}-s will take up its effect.

The final aspect of the utility function specification that is left unresolved is the treatment of individual heterogeneity. So far it was only assumed that individual specific parameters express taste differences across consumers. Nevertheless, an estimable empirical specification requires either an assumption about the theoretical distribution of these parameters, or has to use a proxy for them based on observable individual attributes. In this paper the latter strategy is used. Denote the vector of individual i's attributes by $\mathbf{z}_i = [z_{1i}, z_{2i}, ..., z_{Ri}]$, then the individual specific parameter θ_{1i} will be modelled as:

$$\theta_{1i} = \sum_{r=1}^{R} \theta_{1r} z_{ri}, \tag{1.12}$$

and all other individual specific parameters in $\boldsymbol{\theta}$ are modelled this way.

Finally, a relatively simple logit specification will be assumed for the random part of the utility:

$$\varepsilon_{ij} \sim \text{iid Type I Extreme value}, \tag{1.13}$$

which assumes that unobserved utility components are uncorrelated across individuals and products. This yields the following choice probabilities:

$$P_{ij} = \Lambda \left(\beta_i \mathbf{x}_j + \theta_{1i} L_j + \theta_{2i} P^L L_j \right), \tag{1.14}$$

where Λ denotes the logit function. This specification allows for observed individual heterogeneity in the utility attached to the eco-label. Therefore, the effect of conditional cooperation can be individual specific as well. Moreover, by introducing product specific constants in the vector of product characteristics ($\mathbf{x}_j$), this model also allows for a simple form of unobserved product heterogeneity.

The main difference between (1.14) and a usual discrete choice demand system is the term $\theta_{2i}P^L L_j$. Aggregating the choice probabilities across individuals and labelled products shows the implications of this additional term for the identification of the demand system:

$$P^L = \sum_j L_j \int \Lambda \left(\beta_i \mathbf{x}_j + \theta_{1i}\delta L_j + \theta_{2i}P^L L_j\right) f\left(\mathbf{z}_i|\mathbf{x}_j\right) d\mathbf{z}_i. \tag{1.15}$$

This expression defines an implicit equilibrium condition for P^L. This equilibrium condition and the individual conditional choice probabilities (1.14) jointly specify the demand system. As (1.15) shows, the equilibrium beliefs (P^L) about the mean choice probability of eco-labeled products is endogenously determined in this demand system. This implies that in the individual choice specification the interaction terms involving P^L will be endogenous. In fact, taking into account conditional cooperation as an explanation for the private provision of a public good transforms the logit demand system into a discrete game: individual choices will depend on the equilibrium belief about the fraction of eco-labeled product in the market.

1.3. Identification

There is a well developed body of literature on the identification of discrete choice models like the one described by equations (1.14) and (1.15). In the literature on peer effects, Manski (1993) and Brock and Durlauf (2001, 2003) discuss identification of choice models with strategic interactions. In the empirical industrial organization literature, Bajari, Hong, Krainer, and Nekipelov (2010), Seim (2006) and Sweeting (2009) discuss identification of static games and provide examples for estimation. There is also a large literature on the identification of dynamic games, examples include Rust (1994), Aguirregabiria and Mira (2007), Pesendorfer and Schmidt-Dengler (2003). Bajari, Hong and Nekipelov (2010) provide a recent review of the literature.

The demand system in the current paper is most closely related to the multinomial choice model with strategic interactions of Brock and Durlauf (2003) and the static game studied by Bajari, Hong, Krainer, and Nekipelov (2010). It incorporates a relatively simple form of strategic interactions, because in (1.14) the equilibrium beliefs are homogenous (P^L) and only their contribution to the individual utility varies across individuals (θ_{2i}). However, the choice situation described in the previous section is observed repeatedly in the data. Therefore, the individual choice probabilities in (1.14) are modified to take into account this time series aspect:

$$P_{ijt} = \Lambda\left(\theta_{1i}L_{jt} + \theta_{2i}P_t^L L_{jt} + \sum_{k=1}^{K}\sum_{r=1}^{R}\beta_{kr}z_{rit}x_{kjt}\right), \tag{1.16}$$

where t indexes time. As a consequence, the expected market share of eco-labeled products (P_t^L) will be also time specific.

Assuming that one observes variation in individual choices and average choices, the identification problem can be summarized as follows. Can one conclude that variation in individual choices is due to variation in the average choice or variation in the expected choice is simply a consequence of the variation in individual choices? As the question already suggests, identification is possible if there are covariates that shift the equilibrium belief, but do not influence the individual choices directly (only through the equilibrium). This implies exclusion restrictions on the individual choices: some information about the choice situation should only enter through the equilibrium term. One example for such an exclusion restriction is that average information about observables do not enter individual choices directly.

This identification problem was first described by Manski (1993) in the context of social interactions, who referred to it as the reflection problem. He pointed out that there are three explanations for similar average behavior, which in the consumer choice context are the following:

(1) *Endogenous effects* that are based on expected behavior of other consumers. In case of the choice model described by (1.16): $\theta_{2i} P_t^L L_{jt}$.

(2) *Contextual effects* that capture the influence of different average individual attributes. Such effects are not included in the choice model described by (1.16). Examples for these include: interactions of mean individual attributes and product characteristics: $\sum_{k=1}^{K} \sum_{r=1}^{R} \widetilde{\beta}_{kr} E\left(z_{rit} | \mathbf{x}_{jt}\right) x_{kjt}$.

(3) *Correlational effects* that are observed or unobserved effects common to all individuals at time t. Examples for this are: the average effect of the label dummy ($\theta_{1i} L_{jt}$) or the unconditional mean effects of product characteristics ($\sum_{k=1}^{K} \overline{\beta}_k x_{kjt}$), which are incorporated in (1.16).

As one can see, the choice specification in (1.16) includes both endogenous and correlational effects, but excludes contextual effects. This exclusion restriction is quite standard in demand models and it is reasonable as long as consumers are not willing to pay more for products just because these are more likely to be chosen by certain socioeconomic groups. The assumption that consumption does not signal social status is plausible for some goods (for example toilet paper) and less plausible for others (for example cars). However, even in markets with conspicuous consumption, if there is enough information on individual attributes, it will be plausible to assume that some of these attributes are unrelated to social status (like number of kids) and therefore their average can be excluded from the demand specification.

The exclusion of contextual effects is important for the identification of endogenous effects. This way the equilibrium beliefs (1.15) are conditional on the average individual attributes, but individual choice probabilities are not directly influenced by them (only through the equilibrium beliefs). Therefore shifts in the average attributes shift the equilibrium independently from changes in individual choices. The identification strategy is therefore built on the standard assumption that the consumer's own individual attributes enter their utility functions so that other consumers' attributes do not influence individual choices. This assumption allows one to use moments of the income, schooling etc. distribution to predict expectations about the market share for eco-labeled products at different points in time. Of course, this requires that these distributions vary through time for consumers who buy eco-labeled products.

A simple example gives more insight into the role of excluding contextual effects. Assume that instead of the logit specification in (1.14) a linear probability model is chosen and the eco-label effects are uniform across individuals:

$$\begin{aligned} P_{ijt} &= \theta_1 L_{jt} + \theta_2 P_t^L L_{jt} + \sum_{k=1}^{K} \sum_{r=1}^{R} \beta_{kr} z_{rit} x_{kjt}, \qquad (1.17) \\ P_{ijt} &= y_{ijt} - u_{ijt}, \end{aligned}$$

where the binary variable y_{ijt} indicates that consumer i purchased product j in time t and u_{ijt} is a disturbance. In this setting P_{ijt} is treated as a latent variable. For such a linear model the equilibrium condition (1.15) becomes:

$$P_t^L = \sum_j L_{jt}\left[\int \left(\theta_1 L_{jt} + \theta_2 P_t^L L_{jt} + \sum_{k=1}^{K}\sum_{r=1}^{R}\beta_{kr} z_{rit} x_{kjt} + u_{ijt}\right) f\left(\mathbf{z}_{it}|\mathbf{x}_{jt}\right) d\mathbf{z}_{it}\right],$$

$$P_t^L = \theta_1 \sum_j L_{jt} + \theta_2 P_t^L \sum_j L_{jt} + \sum_j L_{jt} \sum_{k=1}^{K}\sum_{r=1}^{R}\beta_{kr} E\left(z_{rit}|\mathbf{x}_{jt}\right) x_{kjt}, \tag{1.18}$$

$$P_t^L = \frac{\theta_1}{1-\theta_2\sum_j L_{jt}} \sum_j L_{jt} + \frac{\sum_j L_{jt}}{1-\theta_2\sum_j L_{jt}} \sum_{k=1}^{K}\sum_{r=1}^{R}\beta_{kr} \frac{\sum_j L_{jt} E\left(z_{rit}|\mathbf{x}_{jt}\right) x_{kjt}}{\sum_j L_{jt}},$$

where step two assumes $\int u_{ijt} f\left(\mathbf{z}_{it}|\mathbf{x}_{jt}\right) d\mathbf{z}_{it} = 0$. Assuming that $1-\theta_2\sum_j L_{jt} \neq 0$ and plugging (1.18) back to (1.17) yields the following reduced form equation:

$$\begin{aligned} P_{ijt} &= \frac{\theta_1}{1-\theta_2\sum_j L_{jt}} L_{jt} + \\ &+ \left(\frac{\theta_2\sum_j L_{jt}}{1-\theta_2\sum_j L_{jt}} \sum_{k=1}^{K}\sum_{r=1}^{R}\beta_{kr} \frac{\sum_j L_{jt} E\left(z_{rit}|\mathbf{x}_{jt}\right) x_{kjt}}{\sum_j L_{jt}}\right) L_{jt} + \\ &+ \sum_{k=1}^{K}\sum_{r=1}^{R}\beta_{kr} z_{rit} x_{kjt}. \end{aligned} \tag{1.19}$$

(1.19) makes two important points about the identification of θ_1 and θ_2:

(1) θ_2 is identified by the variation in the average interaction terms ($\frac{\sum_j L_{jt} E(z_{rit}|\mathbf{x}_{jt}) x_{kjt}}{\sum_j L_{jt}}$) and the number of labeled products ($\sum_j L_{jt}$). The first requires time variation either in the average individual attributes of consumers who purchase labelled products. Note also that θ_2 enters the regression in a non-linear manner.

(2) θ_1 is not separately identified from θ_2, which is the result of the model structure since θ_1 measures label effects for the hypothetical case when $P_t^L = 0$.

This example highlights that excluding contextual effects - the conditional means of individual attributes - is important to identification, because they capture the same variation in average interaction terms as $\frac{\sum_j L_{jt} E(z_{rit}|\mathbf{x}_{jt}) x_{kjt}}{\sum_j L_{jt}}$. Bajari, Hong, Krainer, and Nekipelov (2010) stress exactly this condition: for some attributes only the individual specific variables should enter the individual utilities, without any average measure based on other consumers' attributes.

Next to the exclusion of contextual effects, the non-linearity of the discrete choice demand system can also be a potential source of identification. As it is originally pointed out by Manski (1993), the non-identification of endogenous effects in a linear regression follows from the linear functional form. Even without the exclusion restrictions discussed above, endogenous effects are identified in non-linear models. Brock and Durlauf (2003) show that in this case sufficient (time) variation in P_t^L identifies endogenous effects. However, such identification always depends on the specific functional form chosen for the choice probabilities. Moreover, in the Blauer Engel case P_t^L varies between 19% and 23% and in this range the logit function is close to linear. Therefore, non-linearity should not be relied upon for identification.

The exclusion of average individual attributes uses the time series dimension of the data to identify the conditional cooperation effect. Unfortunately, in the Blauer Engel dataset there is choice set variation through time: more and more products are eco-labeled. Such choice set variation raises two issues for identification. First, increasing the number of eco-labeled products in the choice set creates a spurious relationship between average choices and individual choices. Even if consumers choose purely randomly the individual will choose an eco-labeled product with higher probability simply because there are more eco-labeled product. On aggregate this translates to an increasing market-share of eco-labeled products and therefore one will find a positive relationship between the two. This way the choice set variation works as a classical confounding factor. Second, choice set variation is endogenous. Firms decide whether to adopt the eco-label based on their expectations about the demand for eco-labeled products.

Both problems suggest that time series variation in the market share of eco-labeled products should not be used to identify the effect of the equilibrium beliefs. Therefore, I will use geographic heterogeneity in individual attributes for identification instead of the time series dimension. Mor precisely, the difference of the geographic unit from the overall trend will be used. Geographic units are defined by the 41 German regions, thus expectations about the public good provision will be region specific: P^L_{th} where h indexes the regions. This amounts to assuming that consumers care about the region specific total amount of the public good rather than the national one. Identification will then be based on the difference $P^L_{th} - P^L_t$. In order to compare results, I will also report estimates based on the time series identification.

An alternative way to overcome the endogeneity problem in the system (1.14) and (1.15) is to use the lagged values of expected market share (for example P^L_{t-1}) as an instrument for the equilibrium beliefs in period t . This strategy avoids the problem of simultaneity of individual choices and the equilibrium, by predicting the latter from past values of the observed eco-labeled market share. This type of instrumentation can be combined with both the time series and the cross-sectional identification proposed in the preceding paragraphs.

I will use a two step method to estimate the demand system in (1.16). Two step estimation is frequently applied to estimate strategic interactions of which Bajari, Hong, Krainer and Nekipelov (2010) is an example. In the first step the aim is to give a flexible approximation of the individual choice probabilities that can be aggregated into an estimate of the equilibrium beliefs. For the demand system in (1.16) this amounts to estimate P_t^L (or P_{th}^L if also regional variation is allowed) in a flexible way. In principle this could be done by estimating the logit demand model with all covariates and a flexible basis function, or even by a linear probability model if covariates are discrete.[2] However, the logit specification simplifies the first step condsiderably in the present case, because it effectively matches the predicted averages individual probabilities to the observed market shares if product specific dummies are included in the estimation. Therefore, the market share for the eco-labeled products can be used as an estimate for the equilibrium beliefs in this case. If regional variation in the expectations is allowed, then the regional market shares are used.

The second step plugs in these estimated beliefs into the structural equation in order to estimate the structural parameters by pseudo maximum likelihood. If the first step is consistent, then the second step will also yield consistent estimates. Standard errors for the strategic interaction effect will have to be bootstrapped generally. However, given the large sample size in the present analysis (340000 observed choices, which accounting for the choice set imply around 17000000 binary decisions), I will continue to use the asymptotic standard errors of the standard logit.

This two step method is facilitated by two assumptions. The first is that the random component of the utility function (ε_{ij}) is private information. This is a quite standard and plausible assumption for a discrete choice demand model. The second is that only one equilibrium is observed in the data., which rules out multiple

[2]In this case the linear probability model estimates the "cells" of the conditional distributions.

equilibria. This is the stronger assumption of the two, given that I observe choices over time and across regions. Period by period estimation (as opposed to pooling all the observations) can make this assumption less strong because it only assumes the poolability of the cross section observations by estimating (1.16) separately for each period.

1.4. Data

The dataset is a unique sample that combines information on toilet paper products labeled with the German eco-label "Blauer Engel" and consumer choice data. The toilet paper market was chosen, because paper products are relatively homogenous, so observable characteristics are likely to capture product heterogeneity well. It is important to control for product heterogeneity in order to measure the label effects in a reliable manner. Moreover, no conspicuous consumption assumption is unlikely to be problematic for these products.

The consumer choice data was provided by GfK Germany, a market research company. It is a household panel dataset, which is a representative sample of German households and contains repeated observations on toilet-paper purchases at a daily frequency for the years 2003-2006. The unit of observation is a daily purchase of a household, and altogether there are 346,320 purchase observations in the dataset.

The consumer choice dataset was in turn merged with data about introduction dates of the national eco-label "Blauer Engel". This allows one to observe the date when a given product was eco-labeled or when the eco-label was removed, making it possible to compare eco-labeled and non-eco-labeled purchases through time and across products. The label information was provided by RAL, the institute responsible for the certification process of these labels.

The combined dataset contains the following product characteristics: purchased quantity, price, shop type, promotion, brand, Blauer Engel label, packaging, color, decoration, layers, fragrance. Table 8 in the Appendix gives a detailed summary of the product characteristics. To draw an aggregate picture of eco-labeled products Table 1 shows the market shares of these products for each sample year. One can observe a slight increase through time, both in the market share and the number of labelled products. This might suggest that consumers value environmental labels and slowly these products gain in market share.

Table 1: Market share of eco-labeled products

	2003	2004	2005	2006
Number of brands	24	29	30	35
Share of eco-labeled products	19.70%	20.20%	21.70%	22.20%
Number of observations	93148	94941	117759	139484

How do eco-labeled products compare to the other products? In order to answer this question, Table 2 compares the modal products with and without eco-label. In terms of observable characteristics, there is almost no difference between eco-labeled and normal products, except for their color and pack size. The important difference is in the prices: the typical eco-labeled product is cheaper than the typical non-eco-labeled product. This might indicate that consumers treat recycled products being of inferior quality, or might simply show the cost advantage of firms from using recycled paper.

Table 2: Average price of modal product in EUR cents/roll*

Year	Without Blauer Engel label		With Blauer Engel
	not recycling no additives white normal price 3 layers no fragrance no decoration tissue 8 roll per pack	not recycling no additives white normal price 3 layers no fragrance no decoration tissue 10 roll per pack	recycling no additives natural normal price 3 layers no fragrance no decoration tissue 8 roll per pack
2003	**31**	24	22
2004	**30**	24	22
2005	28	**23**	19
2006	28	**23**	19

* The modal product changed for non-Blauer Engel labeled products, price of the modal product is indicated by bold.

However, comparing the shop type of the purchase observations, as reported in Table 3, shows that eco-labeled toilet paper was mostly purchased in discount stores. This implies that eco-labeled products compete in a rather different market than the average non-eco-labeled products, where prices are generally lower. Therefore, it is somewhat misleading to compare the modal products simply based on the narrowly defined product characteristics. This cautions against simple price comparisons that do not take into account all the relevant characteristic of the products. The econometric estimation will help to make the proper comparison and provide a ceteris paribus WTP estimate for the Blauer Engel eco-label.

Table 3: Share of shops in Blauer Engel labeled purchases

Shop type	Share
Hypermarket	18
DM	8
Supermarket	5
Discount store	66
Large supermarket	3
Other	0
Department store	0
C&C	1
Internet	0

For the identification strategy to work one also needs sufficient time variation in the individual attributes of consumers who buy eco-labeled products. Table 4 reports the changes in the net earnings distributions for the four sample years. Generally, consumers with higher income have a larger share in all purchases in the last two years. Similar changes can be observed for the other individual attributes as well. This suggests that time variation in these attributes can be indeed used to instrument for the endogenous label effect.

Table 4: Change in the earnings distribution of consumers who purchased eco-labeled products

Net earnings in euro per month	2003	2004	2005	2006
up to 499	4.9	5.5	4.4	4.9
500 - 624	6.4	6.6	6.5	5.2
625 - 749	17.1	15.1	14.3	13.2
750 - 874	9.3	9.5	9.2	8.3
875 - 999	11.7	12.7	12.4	12.0
1000 - 1124	11.5	10.3	10.8	10.7
1125 - 1249	9.9	9.9	10.6	10.1
1250 - 1374	8.9	9.1	8.5	10.3
1375 - 1499	6.2	6.5	6.2	6.7
1500 - 1749	5.3	5.8	6.1	6.1
1750 - 1999	4.3	4.6	5.8	6.7
2000 and more	4.5	4.5	5.3	5.9

1.5. Empirical specification and results

1.5.1. First stage

The aim of the first stage is to give an estimate of the equilibrium beliefs P_t^L (or P_{th}^L if also regional variation is allowed) using all available information. This estimate will be denoted by $\widehat{P}_t^L$ ($\widehat{P}_{th}^L$)and will be used in the second stage to proxy for the equilibrium beliefs. The equilibrium belief is the average conditional probability that a labeled product is chosen at time t in region r:

$$\widehat{P}_t^L = \frac{1}{N_t} \sum_i \sum_j \widehat{p}_{ijt} \left(\mathbf{z}_{it}, \mathbf{x}_{jt}, \boldsymbol{\alpha} \right), \tag{1.20}$$

and

$$\widehat{P}_{th}^L = \frac{1}{N_{th}} \sum_i^{N_{th}} \sum_j \widehat{p}_{ijt} \left(\mathbf{z}_{it}, \mathbf{x}_{jt}, \boldsymbol{\alpha} \right). \tag{1.21}$$

where $\widehat{p}_{ijt}$ denotes the estimated probability that individual i chooses product j at time t conditional on the individual attributes, $\mathbf{z}_{it}$ the characteristics of the product, $\mathbf{x}_{it}$ and the parameter vector $\boldsymbol{\alpha}$. N_t and N_{tr} denote the number of consumers in time t and in region r in time t respectively. The aim of the first stage is to estimate the individual choice probabilities, $\widehat{p}_{ijt}$, in a flexible way, in order to give a consistent estimate of $\widehat{P}_t^L$ and $\widehat{P}_{th}^L$. Given that the individual choice probabilities are specified as a logit model in (1.14), this can be achieved by estimating a logit with a flexible basis function for example. However, the logit specification also offers a much simpler solution to estimating $\widehat{P}_t^L$ and $\widehat{P}_{th}^L$. If the parameter vector contains an alternative specific constant for each period, α_{jt}, then the predicted probabilities from the logit will be 'right' on average: the predicted aggregate logit probabilities will equal the observed market shares. Therefore the conditional predicted market shares will equal the unconditional market shares in case of the logit. This property allows one to use the observed market shares of eco-labeled products as the estimates for $\widehat{P}_t^L$ and $\widehat{P}_{th}^L$. Thus if $y_{ijt} = 1$ if individual i chooses product j at time t and $y_{ijt} = 0$ otherwise then

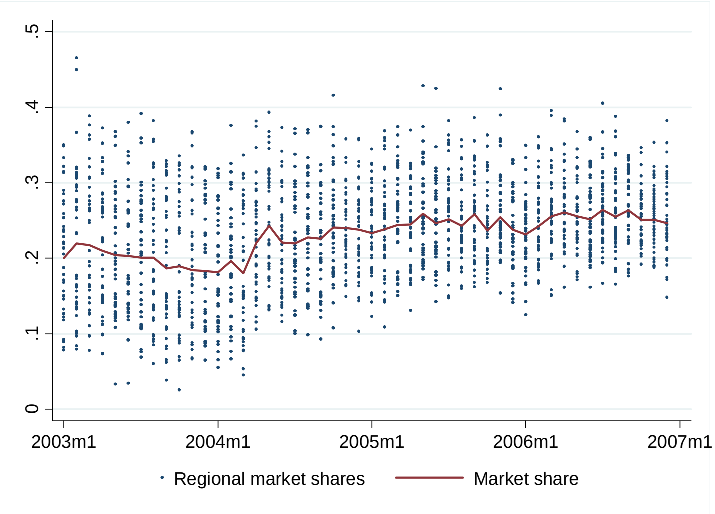
$$\widehat{P}_t^L = \frac{1}{N_t} \sum_i \sum_j y_{ijt} L_{jt}, \tag{1.22}$$

and

$$\widehat{P}^L_{th} = \frac{1}{N_{th}} \sum_{i}^{N_{th}} \sum_{j} y_{ijt} L_{jt}. \tag{1.23}$$

The unit of observation is daily purchase of a household. If a household bought more products at the same time, it will be treated as two separate observations. This simplification assumes that repeated product choices are independent from each other. In order to have a more robust estimate for the equilibrium beliefs, $\widehat{P}^L_t$ and $\widehat{P}^L_{th}$, are calculated as monthly averages.

Graph 1: Market shares of eco-labeled product



Graph 1. shows the evolution of the eco-labeled market share and its regional variation. The graph demonstrates that there is significant variation in predicted probabilities both through time and across regions. Regional variation, expressed in standard deviation, is almost three times higher than time series variation. The ranges also differ: $\widehat{P}^L_{th}$ varies from 2% to 47%, while $\widehat{P}^L_t$ from 18% to 26%. The shift in the predicted probability in the beginning of 2004 reflects the adoption of the Blauer Engel label by a large retailer that sells products under its own brand.

The baseline specification will use $\widehat{P}^L_{th}$ and $\widehat{P}^L_t$ to capture equilibrium beliefs. In addition, however, I will also instrument these probabilities with their lagged values

$\widehat{P}^L_{t-1,h}$ and $\widehat{P}^L_{t-1}$. Lagged probabilities are pre-determined in time t and therefore offer an alternative solution to the endogeneity problem.

1.5.2. Second stage

In the second stage the goal is to estimate the individual demand functions in order to identify the different label effects. The specification is based on (1.14) and includes two types of endogenous effects: the average market share for each period, $\widehat{P}^L_t$ and a time demeaned regional market share, $P^L_{th} - P^L_t$. Products (indexed by j) are defined by their name and the type of the shop where they were bought. Therefore shop choice decisions are implicitly incorporated in the demand model. As usual, the choice set available to different consumers is not observed and it is defined as the products observed in a given month in a given geographical area. Furthermore, individual heterogeneity is captured by the variation in net income categories (z_{irt}, where r indexes the categories in this case). In principle, the second stage should control for the same type of unobserved product heterogeneity as the first stage through introducing alternative specific dummies. However, due to the large dataset and the large number of observed alternatives (approximately 800) this is not feasible. Thus the individual choice specification takes the following form:

$$P_{ijt} = \Lambda \left(\begin{array}{c} \sum_{r=1}^{R} \theta_{1r} z_{irt} L_{jt} + \sum_{r=1}^{R} \theta_{2r} z_{irt} \widehat{P}^L_t L_{jt} + \\ + \sum_{r=1}^{R} \theta_{3r} z_{irt} \left(\widehat{P}^L_{th} - \widehat{P}^L_t \right) L_{jt} + \sum_{k=1}^{K} \sum_{r=1}^{R} \beta_k z_{irt} x_{kjt} \end{array} \right). \tag{1.24}$$

The parameters of interest are θ_{1r}, θ_{2r} and θ_{3r}. The θ_{1r}-s are the parameters on the simple eco-label dummy and correspond to the hypothetical situation when nobody except consumer i is expected to buy an eco-labeled product. Since this situation is not observed it is an out of sample prediction of the model. θ_{2r}-s identify the endogenous eco-label effect based on time variation. As discussed in the section on identification this can be problematic, because of the choice set variation and the endogenous adoption of eco-labels by firms. Therefore θ_{3r}-s identify the endogenous eco-label effect purely from cross-sectional variation, based on the zero mean $\widehat{P}_{th}^{L} - \widehat{P}_{t}^{L}$ variable.

Table 5 reports the mean parameter estimates of the label effects ($\widehat{\theta}_1$, $\widehat{\theta}_2$ and $\widehat{\theta}_3$). Also the coefficients for price per roll and the number of layers variables for comparison. Both the baseline estimates and the time lag instrumented estimates are reported for the endogenous eco-label effects. The two set of estimates are not significantly different from each other. The estimation includes other product characteristics as well to control for additional product heterogeneity: decoration and price type.

Table 5: Parameter estimates

Variable	Baseline	Lagged market share instruments
Blauer Engel dummy (θ1)	-1.223	-1.078
	0.049	0.048
Endogenous effects		
Time variation (θ2)	5.219	4.642
	0.208	0.203
Cross sectional variation (θ3)	3.227	3.361
	0.081	0.100
Price	-0.033	-0.033
	0.000	0.000
Layers	0.273	0.273
	0.004	0.004
Other controls for product heterogeneity	decoration, price type	

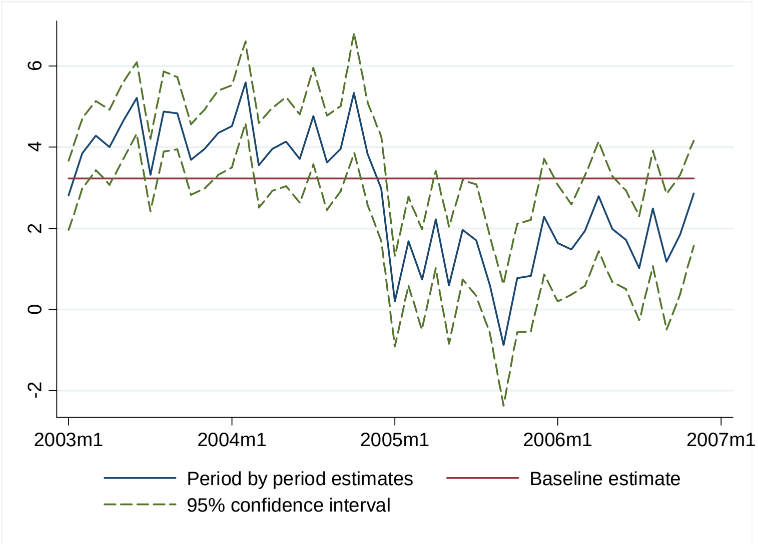
The coefficient of the Blauer Engel dummy ($\widehat{\theta}_1$) has a negative sign. This means that if the consumer thinks that he is the only one buying such products then his demand for an eco-labeled product is actually less than for non-eco-labeled products. This estimate corresponds to the 'warm glow' and pure altruism explanations for public good contribution. Although these explanations cannot be separated within the estimation framework the large negative coefficient implies that their combined effect is negative on consumer demand. The magnitude of the estimated parameter has to be treated cautiously, nevertheless, because in the sample there are no periods with zero share of eco-labeled products.

The estimates of the endogenous eco-label effect ($\widehat{\theta}_2$ and $\widehat{\theta}_3$) have a positive sign and are statistically significant confirming the hypothesis that conditional cooperation is an important explanation for the private provision of a public good. In both specifications the parameter identified from the time variation ($\widehat{\theta}_2$) is higher than the one identified from the cross-sectional variation ($\widehat{\theta}_3$). Endogenous eco-label effects are important, since they change the coefficient of the eco-label dummy considerably. Taking $\widehat{P}^L_{th} - \widehat{P}^L_t$ at its 10th percentile gives a coefficient of -0.25 for the endogenous effect in the baseline specification, while at the 90th it gives 0.28. These are of the same magnitude as the parameter on the number of layers. Similarly, adding $\widehat{\theta}_1$ and $\widehat{\theta}_2$ at the 10th and 90th percentile of $\widehat{P}^L_t$ yields a combined coefficient of -0.23 and 0.13.

It is important to stress that the interpretation of θ_1 is different compared to a specification without endogenous effects. Without endogenous effects the Blauer Engel dummy captures the average effect of the eco-label on consumer choice. Naturally, this average effect can be calculated also from the endogenous-effect specification by combining the coefficients of the Blauer Engel dummy and the endogenous effects (taken at its average 0.243). For the baseline specification this average value is approximately 0.046. Comparing this average coefficient to the one on the number of layers shows that, on average, the Blauer Engel label has virtually no effect on demand compared to this product characteristic.

Specification (1.24) assumes that in each period the same equilibrium is played, because it restricts parameters to be homogenous through time. It is possible to argue, however, that each period a different equilibrium is played with different set of parameters. This implies a period by period estimation of (1.24). In this case endogenous effect is identified automatically from the cross-sectional variation since the data is not pooled across time periods.

Graph 2: Period by period parameter estimates of the endogenous label effect



Graph 2 reports the estimated parameters for the endogenous eco-label effect that were estimated separately for each month. Clearly, there is a regime change after 2004, with parameters higher estimates in the first period and lower in the second. The average parameter in the first regime is 4.0, while it is only 1.5 in the second regime. This suggests that a different equilibrium is realized in these two periods. This change has to be taken into account when evaluating the willingness to pay for the eco-label.

Although the parameter estimates already show the signs and magnitudes of the eco-label effect, on their own they are not very informative for firms or policy-makers. In order to express the effect of the Blauer Engel label in monetary terms one can translate the parameters to willingness to pay (WTP) terms. As Bjørner, Hansen and Russell (2004) show, WTP in logit models can be calculated based on the ratio of the eco-label parameters and the price parameter. I will calculate three WTP measures. The first is the gross WTP, which is based on all three eco-label estimates:

$$WTP(Gross) = -\frac{\widehat{\theta}_1 + \widehat{\theta}_2 \widehat{P}_t^L + \widehat{\theta}_3 \left(\widehat{P}_{th}^L - \widehat{P}_t^L \right)}{\widehat{\beta}^{price}}, \tag{1.25}$$

the second is the WTP based on the endogenous effect identified from time variation:

$$WTP(Time\ variation) = -\frac{\widehat{\theta}_1 + \widehat{\theta}_2 \widehat{P}_t^L}{\widehat{\beta}^{price}}, \tag{1.26}$$

and the third is based on the on the endogenous effect identified from Regional heterogeneity:

$$WTP(Regional\ heterogeneity) = -\frac{\widehat{\theta}_3 \left(\widehat{P}_{th}^L - \widehat{P}_t^L \right)}{\widehat{\beta}^{price}}. \tag{1.27}$$

The second and the third WTP measures basically decompose the gross measure.

Table 6. presents these WTP measures as percentage of the average price for three possible values of the predicted label choice probabilities: 10th and 90th percentiles and means of the predicted equilibrium beliefs. Since $\left(\widehat{P}_{th}^L - \widehat{P}_t^L \right)$ is defined in a way that its mean is zero there is no mean WTP calculated for this case. This property also implies that the mean values are the same for the other two WTP measures. The results are shown for both the baseline and the lag-instrumented specification. The two set of estimates are very close to each other. Under the "Regional heterogeneity" heading the period by period estimates are presented in

the last three columns. These are the averages of the period by period estimates: the overall average and the averages for the two regimes indicated by Graph 2.

Table 6:. Willingness to pay for environmental labels (% of average price per roll)

Value of endogenous effect	Gross		Time variation		Regional heterogeneity				
	Baseline	Lagged IV	Baseline	Lagged IV	Baseline	Lagged IV	Period by period average	Regime 1	Regime2
10th percentile	-57%	-54%	-27%	-23%	-29%	-30%	-42%	-69%	-12%
s.e.	2%	2%	1%	1%	1%	1%	.	.	.
Mean	-1.6%	-0.3%	-1.6%	-0.3%	.	.	-1.4%	-3.5%	1.0%
s.e.	0.7%	0.7%	0.7%	0.7%	.	.	.	.	.
90th percentile	48%	49%	15%	15%	33%	34%	35%	54%	13%
s.e.	1%	1%	1%	1%	1%	1%	.	.	.

First, I discuss the pooled estimates. All three WTP measures indicate that in an environment where consumers' expectations are low (lower 10th percentile of the predicted probabilities) their WTP for eco-labeled products is actually negative. The values range from -23% to 57% of the average price, which is a significant discount. The regional heterogeneity based estimates imply a slightly larger discount. On the contrary, in a high expectation environment (90th percentile of the predicted probabilities) the WTP turns positive due to the endogenous effects and reaches 15% in case of the time variation estimates and 33% in case of the regional heterogeneity estimates. On average the WTP in the sample is essentially zero: at the average value of the predicted probabilities it is only -1.6% of the average price in the baseline specification. Finally, it is also possible to calculate the hypothetical WTP for the simple eco-label dummy estimate $\widehat{\theta}_1$. This value is -150%, which indicates a huge discount required by consumers if they don't expect other consumers to buy the eco-labeled products. Although this result should be taken with a grain of salt it suggests that pure altruism and warm glow effects are unlikely to create significant demand for eco-labeled products.

On average, the period by period estimates show similar WTP values as the pooled, except for indicating a larger discount for low expectation environments. This suggest that pooling does provide a reasonable approximation about the average effects. The two regimes are characterized by sharply differing endogenous effects, nonetheless. In Regime 1 the endogenous effect is roughly twice as large as for the pooled regional heterogeneity estimates. In Regime 2, however, they are only about the third of the pooled estimates. This finding implies that one should not treat endogenous effects as stable through time.

1.6. Conclusion

I argued in this paper that markets with eco-labeled products are examples for the private provision of impure public goods since they also offer the consumer to contribute to a public good, reduction in environmental pollution, next to the usual product characteristics. Therefore these markets can provide market evidence for public good contributions. If one observes consumer choices in such markets one can estimate a discrete choice demand system and provide an estimate for the influence of eco-labels on individual choices. This paper presents a unique dataset containing information about toilet paper products with the German eco-label, the Blauer Engel, and consumer purchases for a representative sample of German households.

However, the public good aspect means that a model of consumer choice in these markets has to incorporate standard explanations for public good provision: pure altruism, warm low and conditional cooperation. It is shown that conditional cooperation implies interdependent consumer choices: a consumer's willingness to contribute to reduced environmental pollution depends on the expected behavior of other consumers. This transforms the usual discrete choice demand system into a game where individual demands are dependent on an aggregate equilibrium condition that defines the expected aggregate probability of eco-labeled purchases. Identification of this modified demand system with endogenous effects is based on the inherent non-linearity of the discrete choice model and the assumption that the consumer does not care who else is buying eco-labeled products and are only concerned about the fraction of the population that buys such products.

The results show that in the sample period the average effect of the Blauer Engel label was practically zero. This questions the role of eco-labels as effective instruments for voluntary reduction in environmental pollution. However, the results also point out that endogenous effects alter the impact of the Blauer Engel label significantly providing evidence that conditional cooperation is important in markets with eco-labels. Consequently focusing exclusively on average effects can be misleading in evaluating eco-labels. This result seems to be robust across different specifications. It also implies, however, that WTP for the Blauer Engel label can be negative as well. Moreover, the period by period estimates suggest that the endogenous eco-label effect is not stable through time.

From the perspective of environmental policy the significance of endogenous effects is an ambiguous finding. It suggests that eco-labels can indeed influence demand, but only if there is already a 'critical mass' of consumers buying such products. As a consequence small scale adoption of eco-labels by producers is unlikely to be effective. Finally, regime changes in the endogenous effect suggest that eco-labels can not be regarded as a robust policy instrument.

1.7. Appendix

Table 7: Descriptive statistics of product characteristics

Variable	Varieties	Frequency	%
Packaging	plastic	437,852	99.04
	paper	4,262	0.96
Decoration	with	67,588	15.31
	without	373,798	84.69
Fragrance	with	13,517	3.1
	without	422,166	96.9
Private label	yes	209,709	47.09
	no	235,623	52.91
Layers	1	6,653	1.5
	2	32,576	7.3
	3	308,697	69.3
	4	96,038	21.6
	5	1,364	0.3
Type	tissue	439,176	98.9
	crepe	4,889	1.1
Additive	none	428,993	96.81
	chamomile	12,079	2.73
	aloe vera	2,043	0.46
Color	white	311,811	71.61
	natural	89,958	20.66
	colored	33,198	7.62
	other	445	0.1
Pack size	1	415	0.1
	2	6,084	1.4
	4	1,388	0.3
	6	15,200	3.4
	8	205,859	46.2
	9	2,637	0.6
	10	164,342	36.9
	11	5	0
	12	23,366	5.3
	15	72	0
	16	15,848	3.6
	18	2,335	0.5
	20	2,223	0.5
	23	433	0.1
	24	5,087	1.1
	30	31	0
	64	7	0

Table 8: Descriptive statistics of demographical variables

Variable	Classification	Frequency	%
Age of the primary earner	<19	218	0.05
	20-24	7,620	1.71
	25-29	24,413	5.48
	30-34	42,654	9.58
	35-39	60,865	13.67
	40-44	60,622	13.62
	45-49	52,755	11.85
	50-54	46,948	10.55
	55-59	34,960	7.85
	60-64	34,916	7.84
	65-69	34,367	7.72
	70-74	25,051	5.63
	75<	19,803	4.45
Occupation	pensioner	108,479	24.36
	skilled manual	58,289	13.09
	low/middle civil servant	15,068	3.38
	qualified clerical worker	115,147	25.86
	clerical worker	26,692	5.99
	unemployed	19,822	4.45
	leading civil servant	5,206	1.17
	gehobene beamte	12,666	2.84
	housewife	4,002	0.9
	manager	25,699	5.77
	highly skilled manual	4,509	1.01
	unskilled manual	16,134	3.62
	self-employed	17,879	4.01
	professional	3,702	0.83
	farmer	667	0.15
	student	6,014	1.35
	widow/widower non-working	2,366	0.53
	unpaid leave	1,888	0.42
	apprentice	1,103	0.25
Net income per person in EUR	<499	14,489	3.25
	500 - 624	20,005	4.49
	625 - 749	53,903	12.1
	750 - 874	34,756	7.8
	875 - 999	52,857	11.87
	1000 - 1124	48,943	10.99
	1125 - 1249	46,837	10.52
	1250 - 1374	47,927	10.76
	1375 - 1499	34,259	7.69
	1500 - 1749	29,118	6.54
	1750 - 1999	32,168	7.22
	>2000	30,070	6.75
Education of the primary earner	elementary school without apprenticeship	15,469	3.47
	elementary school with apprenticeship	119,173	26.76
	vocational training school	6,767	1.52
	vocational training school with apprenticeship	113,501	25.49
	vocational secondary school	55,430	12.45
	G.C.E. without occupational training	5,818	1.31
	G.C.E. with occupational training	27,073	6.08
	college/university	102,099	22.93

CHAPTER 2

Not an average story: Asymmetric price transmission in the Hungarian gasoline retail market

2.1. Introduction

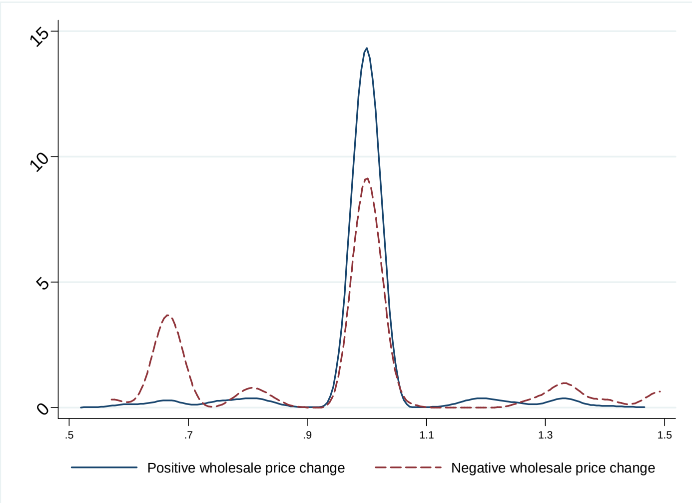
Gasoline is one of the basic commodities in modern economies. Therefore, price movements in this market attract the attention of both consumers and policymakers. One common concern is the asymmetric retail price response to wholesale price changes. In this case, gasoline retail prices rise faster after a cost increase than they decline after a cost reduction. Such an asymmetric retail price response implies that consumers are saving less on a price drop than their additional costs are due to an equivalent price increase. Since gasoline expenditure is a sizeable fraction of households' budget, they can be sensitive to such asymmetries. From a competition policy perspective it is important to determine whether the asymmetric retail price response is a result of anti-competitive behavior of firms or it can be explained by firms' costs and consumer behavior.[1]

Empirical evidence on pricing asymmetries based on cross-sectional information about firms and markets is relatively scarce. The current paper contributes to this type of disaggregated evidence. It studies how station-level retail prices respond to wholesale price changes in the Hungarian gasoline market. To have a first quick

[1]Besides the gasoline market, asymmetric price transmission was studied in a wide range of other markets and at the aggregate level as well. For example Meyer and Cramon-Taubdel (2004) review the literature focusing on the agricultural sector, while Peltzman (2000) and Álvarez et al (2006) provide evidence for asymmetric price adjustment across various products and sectors. For Hungary Reiff and Karadi (2007) provide evidence of asymmetric inflation response to VAT changes.

look at the possible pricing asymmetry, Graph 1 shows the ratio of the retail and the wholesale price change. The distribution of this ratio is depicted separately for increasing and decreasing wholesale price periods.[2] Both distributions are centered around one, though the distribution of the negative wholesale price changes is bi-modal and there is significant mass concentrated around 0.7. This suggest that a sizeable fraction of the gasoline stations set prices asymmetrically, though no aggregate asymmetry is expected. These two stylized facts will serve as the null hypothesis of the paper.

Graph 1: Ratio of retail and wholesale price change, distribution over all observations



In order to determine what explains asymmetric pricing an error correction model is estimated with different set of parameters for increasing and decreasing wholesale price periods. The station level data allows to explore how the pricing asymmetry depends on station specific observables (location, brand, number of competitors,

[2]For ease of comparison, the price changes for the decreasing wholesale price periods are multiplied by minus one. A value below unity indicates that the retail prices changes less than the wholesale price change.

type of competitors). Linking station characteristics and pricing behavior can also shed light on the possible explanations behind pricing asymmetries.

Explanations for the pricing asymmetry in gasoline retail markets fall into four categories: tacit collusion models, inventory or related adjustment cost models, search models with frictions and price cycle models. The first explanation was traditionally the main interest of the literature, since in this case, asymmetries might signal collusive behavior. As summarized by Borenstein and Shepard (1996), collusion is usually modelled in a trigger strategy framework, where firms charge a collusive price until one of them deviates, which starts an undercutting phase towards the Bertrand price equilibrium. Negative wholesale price changes can start an undercutting phase, but due to collusion, firms wait for some periods before deviation begins. Wholesale price increases, however, do not have such an effect on collusion if margins are small enough for cost increases to bite instantaneously. Therefore, collusion based on a trigger strategy can lead to an asymmetric retail price response.

The inventory cost based argument for price asymmetry is based on the nonnegativity constraint on inventories. Nonnegative inventories imply that the cost of decreasing inventories has to increase sharply below some threshold, which is not necessarily the case for increasing inventories. This inventory cost asymmetry can then translate to asymmetry in retail prices. Examples for such models include Reagan and Weitzman (1982) or Borenstein and Shepard (2002).

Search models also offer an explanation for pricing asymmetries. Costly search might prevent consumers from making optimal decisions. Therefore, they might choose higher prices even if costs decreased and are likely to accept higher prices more easily when costs increase. Alternatively, as Lewis (2011) explains, search frictions can also be generated by consumers who base their decisions on reference prices and therefore are less price sensitive for downward price movements.

Finally, price cycles - even if unrelated to wholesale changes - can also explain the evidence on asymmetric transmission. Following Maskin and Tirole (1988), price cycles are modelled by firms that set their price above marginal costs and then gradually decrease it to a point when they restart the cycle even at the cost of current period losses rather than to continue the undercutting phase. This data-generating process creates more observations with decreasing prices in smaller increments and fewer increases in larger increments. Therefore, downward retail price changes will have a weaker correlation with cost shocks that are otherwise unrelated to the price cycle. As Eckert (2002) points out, this behavior can produce asymmetric dynamics that is nevertheless unrelated to cost shocks. Furthermore, price cycles can also interact with cost related price asymmetry. Lewis and Noel (2011) find that the pass-through of cost changes to retail prices is 2 to 3 times faster in cycling markets than in non-cycling markets. This implies that markets in which there are frequent, non-cost related retail price changes, cost related price changes happen faster.

Geweke (2004) stresses that there is no reliable identification strategy to differentiate among these explanations. Indeed,the central reference of the literature, Borenstein, Cameron and Gilbert (1997, hereafter BCG), finds evidence for collusion, inventory constraint and search-cost explanations as well. Similarly, Lewis (2011) provides evidence for a reference price search model and Eckert (2002) for alternating pricing regimes that generate price cycles. The main difficulty in identification is that the same station-specific observables can be related to more than one theoretical explanation. Moreover, it is unclear whether the observables capture station specific cost heterogeneity or they truly identify proposed explanations behind the asymmetric pricing patterns.

The results of this paper based on the Hungarian sample confirm the stylized fact displayed by the retail-wholesale price change ratios in Graph 1. The estimates show that although retail price changes are almost symmetric on aggregate, there is a subset of stations that follow an asymmetric pricing strategy. On average, these stations earn 0.5 HUF on each 1 HUF wholesale price increase compared to a wholesale price reduction in the first three weeks following the wholesale price change. While such a difference is substantial when compared to the average markup (above 10%), it is negligible for consumers. Having a closer look at station characteristics reveals that asymmetric pricing is a brand property: there are four brands that change prices with sizeable asymmetry. The brands that price asymmetrically have small market share (below 10%) and are not vertically integrated. Other observables, like the number of competitors or types of competitors do not explain asymmetric retail price response.

These results imply that in the same local market there are firms that price symmetrically and firms that price asymmetrically. This finding does not support collusion and search based explanations that depend on market level interactions among firms and consumers. Instead it points towards the role of adjustment costs as an explanation for asymmetric retail price response. Moreover, the result that the number of competitors and type of competitors does not explain asymmetric retail price responses lends additional support to the claim that pricing asymmetry does not necessarily imply collusive behavior. Finally, regarding price cycles, the

Hungarian dataset is not suitable to make any conclusions, because the data is not recorded at high enough frequency.

An important and unique feature of the Hungarian data is that the dominant wholesaler in the market leaks each week by how much it will change the list wholesale price. This public wholesale price information has three major implications. First of all, it allows one to observe the actual cost shock to gasoline stations. Other European studies usually proxy wholesale prices by the Rotterdam spot market price, which can be a quite rough approximation to the actual cost shock of gasoline stations. For example, in the Hungarian market the list price follows the market price with considerable lag in most periods (as shown by Graph 10 in the Appendix). Second, the fact that wholesale price changes are announced at a regular weekly schedule ensures that the wholesale price changes are observed at the same frequency as retail prices. Third, this public price information suggests that competitors and consumers are well informed in this market, which is consistent with finding no retail price asymmetry.

The paper is organized as follows. First, the findings of the existing literature is discussed focusing on explanations for asymmetric price transmission and estimation methods. Second, the market structure and the pricing behavior in the Hungarian gasoline market is described and the dataset is introduced. Third, the empirical methodology is discussed focusing on the question how far it can be interpreted as a reduced form price equilibrium. Fourth, the results are presented. Finally, the conclusion discusses their implications for possible theoretical explanations of asymmetric retail price response.

2.2. Evidence on asymmetric price transmission in gasoline retail markets

Generally, three price transmission channels are investigated following BCG (1997):

(1) transmission of crude oil to wholesale prices; examples include Borenstein and Shepard (2002), Bachmeier and Griffin (2003);

(2) transmission of wholesale to retail prices; for example Eckert (2002), Deltas (2008), Lewis and Noel (2010), Asplund, Erikson and Freiberg (2000);

(3) transmission of crude oil to retail prices, like in Radchenko (2005).

The present paper focuses exclusively on the second channel, the wholesale-retail price transmission.[3] The third transmission channel contains the first two channels and therefore it is unclear whether transmission asymmetry along this channel derives from the sluggish response of wholesale or retail prices. In order to differentiate between these two possibilities, one has to focus on the first two channels, which assume that the researcher is able to obtain a good measure of wholesale prices. For this, many studies for the US use unbranded rack prices that are prices of unbranded gasoline at the local distribution terminal. Alternative measures are dealer tank wagon prices and bulk spot market prices. Starting form Bacon (1991), European studies, like Grasso and Manera (2007), use the Rotterdam spot market price as a wholesale price measure. Although retail prices are in general easily observable, there are examples for using list or recommended retail prices instead of observed ones, for example in Asplund, Erikson and Freiberg (2000) and Bettendorf, van der Gees and Varkevisser (2003).

[3] As Graph 11 in the appendix suggests that there is some indication of wholesale price smoothing by the largest retailer, MOL, when compared to the oil price (Ural (Med)).

Traditionally, asymmetric price transmission in gasoline markets was studied in a time series framework. In these studies the question is how the average price in a market responds to cost changes. Most studies report slower downward price adjustment. As reviewed by Frey and Manera (2007), the most commonly applied methodology is some variant of an error correction model (ECM) with retail and wholesale prices forming a long-run equilibrium relationship. These models are sometimes generalized to multi-equation vector auto regression or include regime switching as well. Estimation methods for the ECM vary from Engel-Granger two-step, through Stock and Watson (1993) to Johanssen's maximum likelihood method. The frequency of the series ranges from monthly to daily, though many studies (for example Asplund, Erikson and Freiberg, 2000) point out that observational frequency should correspond to the wholesale price setting frequency in the market. Even the choice of the day when prices are observed may not be innocuous as Bettendorf, van der Gees and Varkevisser (2003) demonstrate.

Recently, a series of papers analyze firm level panel data instead of single time series: Verlinda (2008), Faber (2009) and to some extent Lewis (2011). These studies concentrate on the price response of firms to cost changes. Contrary to time series studies such disaggregated perspective allows one to estimate not only the average effects.but also how the asymmetry depends on observable features of stations or local markets. These studies found differences in the retail price transmission according to market and station characteristics. Using data from Southern California, Verlinda (2008) finds that transmission asymmetry depends on the brand of the station, proximity of competing stations and local market features. Another example is Deltas (2008), who shows that transmission dynamics depends on margins, an indicator for the intensity of competition.On the contrary, using Dutch data, Faber

(2009) finds that asymmetry is unrelated to station characteristics. Although Lewis (2011) uses a panel dataset as well, he estimates the price-cost equation on a time series and uses the panel dimension only to estimate the correspondence between price dispersion and margin or prices.

Another advantage of station level data is that it corresponds better to theoretical objects of interest. The theoretical models connect prices of a firm to its costs and market specific variables like the number of competitors. Firms optimize based on these variables. Station level data allows to observe exactly these objects, while aggregate data allows to observe only average variables that are not the outcome of any optimization process. Hence it is more plausible to interpret the long run price equation in the ECM as a reduced form price equilibrium. In the section on empirical specification I discuss in detail how the ECM specification corresponds to different models of price equilibrium.

Station level data, however, also raises the question of how to incorporate station specific heterogeneity into the traditional ECM framework. Such heterogeneity can enter all three components of the ECM:

(1) the equilibrium pricing equation: the equilibrium price can be firm or station specific, due to differences in costs and local demand conditions;
(2) the dynamic adjustment model: the response of prices to cost changes may be also station-specific due to varying local competition structures and brand-specific pricing strategies.
(3) the error-correction mechanism: the speed of adjustment towards the equilibrium price might differ across different pricing regimes and stations types.

Verlinda (2008) estimates both a specification with homogenous parameters and one with station-specific random parameters. The latter is estimated in a hierarchical-Bayes framework and takes into consideration both observed and unobserved heterogeneity in all three components of the ECM. He finds that the response to negative shocks is slower: after a negative shock it takes around 6 weeks to make the same adjustment as in the case of positive shocks. Faber (2009) presents both pooled and station by station Engel-Granger two-step estimates, but does not incorporate directly station or market level observables into the estimation. Both estimations yield very similar results, with asymmetry lasting for 2 days on average in both cases, which is very weak evidence for asymmetric transmission. The station by station estimates also reveal that only 38% of the stations price asymmetrically. Lewis (2011) only allows for heterogeneity between high and low mark-up regimes in the error-correction specification by estimating a threshold ECM.

The current paper uses station level data to estimate the pricing asymmetry and uses two different estimators. The first is the mean group estimator, which allows for station-specific parameters, and its treatment of heterogeneity is similar to that of Verlinda (2008). However, this specification assumes away variation in time. Therefore a second estimator is also proposed, where station-specific heterogeneity is modelled by observable characteristics that can vary through time. This second estimator imposes restrictions on cross-station heterogeneity but allows for time variant heterogeneity.

2.3. Market structure and data

In the Hungarian gasoline market vertical integration is a dominant feature since all wholesalers are also active on the retail market. There are eight major brands, the others are treated as white (unbranded) stations during the analysis (and receive brand code 1 in graphs and tables). The wholesale market is extremely concentrated: the largest firm on the wholesale market (MOL) serves approximately 80% of the stations from its two refineries in Százhalombatta (Hungary) and Bratislava (Slovakia) (OECD, 2008), while the second (OMV) and third (Lukoil) largest firms on the wholesale market serve predominantly their own retail chains and control 10% and 5% respectively. Other integrated firms have negligible share in the wholesale market.

Compared to the wholesale market, the retail market is less concentrated, with the four main players (MOL, OMV, Shell and Agip) owning almost 70% of petrol stations. Three other international brands were also present in 2007: Esso (Exxon), Jet (ConocoPhillips) and Lukoil. From the early 2000s onwards supermarket chains (mainly Tesco) began opening discount stations, and there has been a large number of independent retailers (white pumps). Since 2007, the Hungarian retail market saw two mergers: in February 2007 Lukoil took over Jet's and in July 2007 Agip took over Esso's network of stations. Table 1 shows the market shares of the eight largest brands before and after the mergers in the total number of observations. These figures indicate that the market structure did not change significantly besides the mergers. Also the market share of small brands and non-branded 'white' stations is stable around little more than 20% of the market. This part of the market is important since it is expected to include the most competitive firms in terms of prices.

Table 1: Brands: market shares and margins

Brand	Market share 2007	Market share 2008	Average margin 2007-2008
Small or non-branded	20.9	23.1	4.5
Agip	8.4	10.6	5.7
Esso	2.7		3.9
Jet	2.3	5.8	3.6
Lukoil	3.6		3.6
MOL	30.8	28.0	6.1
OMV	10.5	13.0	5.9
Shell	17.1	15.9	5.9
Tesco	3.7	3.6	5.0

Competitors can have a strong influence on price setting. However, there are several ways to define the competitors of a station. I use a geographic market definition based on the statistical municipalities defined by the Hungarian Central Statistical Office (KSH). These geographic units are defined as distinct geographical areas where inhabitants perform the majority of their social and economic activities. Panel A of Table 2 shows that most stations have many (10 or more) competitors and monopoly markets have relatively small share. Looking at Panel B shows that each large brand (MOL, OMV, Shell and the non-branded stations) is a competitor of approximately 60% of the stations. Nonetheless, even the small brands are relatively frequently competitors of other stations. The geographic distribution of stations is fairly even across counties as shown in panel C of Table 2. There are only two concentrated areas: the capital (county code 5) and its surrounding county Pest (county code14).

The sample used in this paper is a unique panel of petrol stations with weekly observations of retail prices, station characteristics and wholesale prices for 68 weeks in 2007 and 2008. In total, there are 83387 observations for 1291 petrol stations, which cover more than 95% of stations in Hungary. The structure of the dataset is hierarchical. First, there is the level of stations that are the basic units of observation. Second, there are local markets, defined by the statistical municipalities. Competitors are defined at this level, and generally market structure variables refer to this level of observation. Finally, there is the country level that is the collection of all stations. At this level brands are the relevant unit of observation both in the retail and at the wholesale market. The source of the retail data is a private gasoline price comparison website - www.holtankoljak.hu (Where To Fill Up?) - that makes price data publicly available. The website collected the data between Wednesday and Friday each week and refreshed the data continuously. The retail price observations in the sample were downloaded each Friday.

Table 2: Distribution of station characteristics

A: Number of competitors		B: Percentage of stations competing with a given brand		C: County			
Competitors	%	Brand	%	County	%	County	%
0	1.3	Small or non-branded	64.0	1	6.4	11	3.6
1	4.8			2	4.2	12	3.4
2	3.7	Agip	56.0	3	3.9	13	1.7
3	6.8	Esso	35.0	4	5.5	14	11.3
4	7.5	Jet	37.9	5	14.3	15	3.5
5	5.7	Lukoil	41.7	6	4.6	16	4.1
6	7.1	MOL	66.4	7	4.5	17	3.2
7	5.8	OMV	57.1	8	4.7	18	4.1
8	3.9	Shell	60.4	9	4.9	19	4.6
9	2.0	Tesco	43.1	10	3.9	20	3.4
10 or more	51.3						

An important feature of the dataset is that the wholesale list prices of the largest wholesaler, MOL, are publicly observable. Each Monday the online media reports the expected list price change of this wholesaler. Changes are effective from 12:00 p.m. that Wednesday and the price comparison website collects information from all stations by Thursday. During the sample period the pre-announced price changes always coincided with actual ones (which are made public as well). The timing of the wholesale price announcements also implies that wholesale price changes are pre-determined with respect to retail prices that are observed each Friday. Given the high market share of this wholesaler in the upstream market (80%), its list wholesale price is a good measure for all retail stations in the market. It is quite telling that the media interprets the wholesale price changes as changes in the price of gasoline as such, implying the same changes to retail prices. Nevertheless, the actual wholesale price paid by retailers may vary due to discounts that are not observed. This has to be controlled in the econometric specification.

Table 3 reports the descriptive statistics for the price variables separately for increasing, decreasing and no change periods. In total there are 25 weeks when wholesale prices increased, 6 weeks when they decreased and 37 weeks when there was no change. Mean retail and wholesale prices differ across these categories. Prices are on average the highest in decreasing periods, followed by no change and increasing periods. These differences are due to the trending nature of the price series: prices grew approximately by 20% in the sample period. Nevertheless, average retail price changes in decreasing and increasing periods are almost the same and the same holds for wholesale price changes as well.

As it was already mentioned in the introduction, the ratio of the retail and wholesale price changes can serve as a simple indicator of relative change. Table 3 shows the average ratio, which is smaller for decreasing periods, but this difference is very small and statistically insignificant. The standard deviation of the average ratio is quite large for both decreasing and increasing periods. This suggests sizeable heterogeneity in the wholesale price transmission. The mark-up can be another indicator of transmission asymmetry, if it is higher in decreasing wholesale price periods. If the mark-up is higher in decreasing periods it can also indicate transmission asymmetry. This is indeed the case here, although once again the variation in these averages is too large and the difference is not significant. Both the retail-wholesale ratio and the mark-up suggest that the average asymmetry is small, but there can be large differences among stations. The empirical analysis tests this hypothesis and connects the variation in the asymmetry to station and market characteristics.

Table 3: Descriptive statistics of price variables

Variable	Direction of wholesale change	Mean	Sd.	Obs.
Retail price				
	Decreasing	297.7	8.6	7,021
	Increasing	286.2	17.0	27,520
	No change	289.2	17.2	41,576
Wholesale price	Decreasing	282.2	7.2	7,658
	Increasing	271.7	15.8	30,376
	No change	274.6	16.0	45,353
Retail price change	Decreasing	-3.6	2.2	6,661
	Increasing	3.8	1.8	26,233
	No change	-0.0	0.9	39,657
Wholesale price change	Decreasing	-3.9	2.1	7,658
	Increasing	3.9	1.3	30,376
	No change	0.0	0.0	45,353
Ratio of changes	Decreasing	0.95	0.39	6,661
	Increasing	0.98	0.37	26,233
	No change	.	.	.
Markup				
	Decreasing	15.4	4.4	7,021
	Increasing	14.7	4.8	27,520
	No change	14.7	4.7	41,576

2.4. Empirical specification

2.4.1. Controlling for station level heterogeneity in the ECM

Following the literature I use an error correction model to estimate the wholesale-retail price transmission asymmetry. The error correction model combines a static price equilibrium and a dynamic adjustment to deviations from this equilibrium. This dynamic response, however, is not explicitly modelled within the price equilibrium. It is rather a statistical relationship that describes the response of retail prices to a change either in the factors that determine the equilibrium (mainly the wholesale price) or to a random shock that moves prices out of equilibrium.

Assuming that the sample is a panel of stations (indexed by $i = 1, .., N$) that are observed through time (indexed by $t = 1, .., T$) with their retail (p_{it}) and uniform wholesale (c_t) prices the standard specification of the equilibrium retail price is:

$$p_{it} = \theta_{it} c_t + \nu_{it}, \tag{2.1}$$

where θ_{it} is a station-time specific parameter that expresses the long term pass-through and ν_{it} is the disturbance. The linear specification allows one to study the price-cost correspondence as a cointegrating relationship in case p_{it} and c_t are non-stationary. *(2.1)* is essentially a reduced form of a first order condition derived from the firm's profit maximization problem and it is supposed to characterize retail prices in equilibrium. Since the Hungarian price data is non-stationary, this price equilibrium is assumed to follow a cointegrating relationship.

The dynamic adjustment describes the relationship between retail and wholesale price changes and therefore it is the primary interest for the estimation of price asymmetries. A characteristic feature of the ECM model is that only shifts in the explanatory variables change the equilibrium relationship, random disturbances are eliminated through the equilibrium adjustment mechanism. For asymmetric price transmission this implies that wholesale and retail prices cannot drift apart even if there is a pricing asymmetry. The ECM is specified as a distributed lag model supplemented by an equilibrium correction term that describes the adjustment to the static price equilibrium in case of a random shock:

$$\Delta p_{it} = \sum_{j=0}^{J} \beta_{itj}^{+} \Delta^{+} c_{t-j} + \sum_{j=0}^{J} \beta_{itj}^{-} \Delta^{-} c_{t-j} + \lambda_{it} \left(p_{it} - \alpha_{it} - \theta_{it} c_t \right) + u_{it}, \tag{2.2}$$

where β_{itj}-s are parameters on contemporaneous and lagged changes of wholesale prices (the number of lags is J), while λ_{it}-s show the speed of adjustment of the equilibrium correction term $(p_{it} - \alpha_{it} - \theta_{it} c_t)$. The '+' ('-') superscript denotes a positive (negative) wholesale price change and its corresponding parameter. This separation of the β_{itj}-s allows to identify the asymmetric price transmission. The disturbance of the dynamic model is u_{it}. Compared to standard specifications *(2.2)* does not contain lagged values of retail prices. These autoregressive terms are omitted here for two reasons. First, in the sample there is almost no change in retail prices except when there is a wholesale price change. Therefore, lagged retail price changes are almost perfectly collinear with lagged wholesale price changes and the two separate sets of parameters would not be possible to identify. Second, given the price setting game observed in this specific market, wholesale price changes are predetermined with respect to retail price changes, so there is no need to include lagged retail price changes on endogeneity grounds.

In *(2.1)* and *(2.2)* the parameters $\left[\theta_{it}, \boldsymbol{\beta}_{it}^{+}, \boldsymbol{\beta}_{it}^{-}, \lambda_{it}\right]$ are described as station and time specific and therefore have to be restricted in order to estimate them using panel data. Three restrictions will be applied, each of which makes different assumptions about the unobserved station level heterogeneity. The first parameter restriction assumes that both the price equilibrium and the dynamic adjustment is the same for all firms:

$$\begin{aligned} \theta_{it} &= \theta, \\ \beta^{+}_{itj} &= \beta^{+}_{j}, \\ \beta^{-}_{itj} &= \beta^{-}_{j}, \\ \lambda_{it} &= \lambda. \end{aligned} \tag{2.3}$$

(2.3) implies homogenous parameters for all stations in both regressions and I will refer to this specification as 'homogenous panel'. In this specification, station level heterogeneity is captured only by the disturbance, which is described as a classical panel fixed effect model for both the price equilibrium and the dynamic adjustment:

$$\begin{aligned} \nu_{it} &= \alpha_i + \epsilon_{it}, \\ u_{it} &= \beta_{0i} + \varepsilon_{it}, \end{aligned} \tag{2.4}$$

where α_i and β_{0i} are fixed effects and ϵ_{it} and ε_{it} are *i.i.d.* random components. This homogenous specification can capture the mean price equilibrium, the average adjustment dynamics and give an estimate of the average asymmetry in the market. The results can be easily compared to those of aggregate time series studies, but they provide no evidence about the differences in the two relationships among stations. This lack of cross-sectional comparison also means that this estimation does not provide any direct information about the source of price asymmetry.

The second restriction assumes that parameters are station-specific in both the price equilibrium and the dynamic adjustment equations:

$$\begin{aligned} \theta_{it} &= \theta_i, \\ \beta^{+}_{itj} &= \beta^{+}_{ij}, \\ \beta^{-}_{itj} &= \beta^{-}_{ij}, \\ \lambda_{it} &= \lambda_i. \end{aligned} \tag{2.5}$$

This specification will be referred to as 'unobserved heterogeneity', because it allows for unobserved heterogeneity in the parameters. Practically this specification implies a station by station estimation, which means that the disturbances are described by the same station specific fixed effect as in (2.4).

The third restriction assumes that heterogeneity in the parameters can be accounted for by the following station level observables:

- brand of the station ($brand$),
- number of competitors ($number\_competitors$),
- competitor brands within the local market ($brand\_competitor$),
- county in which the station is located ($county$).

All these characteristics are categorical variables and therefore their combinations effectively identify different groups of stations. Therefore θ_{it} will be modelled as follows:

$$\theta_{it} = \theta_1 brand_{it} + \theta_2 number\_competitors_{it} + \sum_{m=1}^{8} \theta_{3m} brand\_competitor_{mit} + \theta_4 county_{it}, \tag{2.6}$$

and the $\boldsymbol{\beta}_{it}^{+}$-s, $\boldsymbol{\beta}_{it}^{-}$-s, λ_{it}-s are expressed similarly. This restriction allows for both time and station specific variation in the parameters by pooling observations across stations. In comparison the unobserved heterogeneity specification allows only for station specific variation although along this dimension it is less restrictive than the observed heterogeneity specification because it only uses the time dimension of the data. The observed heterogeneity specification also includes fixed effects as described in *(2.4)*.

Similarly to classical time series studies the homogenous panel specification yields one aggregate (market level) set of parameters. However, the unobserved and observed heterogeneity specifications provide a distribution for each parameter. These distributions provide information not only about the aggregate pricing behavior but also about the importance of observed characteristics in explaining pricing asymmetries. In order to summarize the estimation results, I will present aggregate estimates, conditional parameter distributions and partial effects of observed characteristics. As to the latter, it is straightforward to calculate partial effects for the observed heterogeneity specification and therefore an aggregate estimate can be given as the mean partial effect. In case of the unobserved heterogeneity specification both the aggregate estimates and the partial effects of the characteristics are estimated using a mean group estimator, which is basically an unweighted average of the estimated parameter distribution.

2.4.2. The ECM as a reduced form price equilibrium

Although the ECM as presented in *(2.1)* and *(2.2)* is a reduced form system, it is worthwhile to discuss briefly the interpretation of the equilibrium price equation *(2.1)* in order to clarify how this widespread empirical specification corresponds to a theoretically plausible notion of price equilibrium. The differentiated Bertrand oligopoly is a good starting point for this exercise, since this model captures the pricing behavior of a gasoline station and its local competitors quite well. In the gasoline retail market product differentiation is due to differences in stations' locations and characteristics, like brand, the presence of a shop etc. In the differentiated Bertrand oligopoly station i faces the residual demand $D_{it}\left(p_{it}, p_{-it}\right)$ in time t that depends both on its own price, p_{it} and its local competitors' prices, p_{-it}. If marginal costs of station i are $C'_{it}\left(D_{it}\left(p_{it}, p_{-i}\right)\right)$, then the first order condition in equilibrium $\left(p^*_{it}, p^*_{-it}\right)$ is the following:

$$\begin{aligned} 0 &= \left[p^*_{it} - C'_{it}\left(D_{it}\left(p^*_{it}, p^*_{-it}\right)\right)\right] \frac{\partial D_{it}\left(p^*_{it}, p^*_{-it}\right)}{\partial p_{it}} + D_i\left(p^*_{it}, p^*_{-it}\right) \qquad (2.7) \\ p^*_{it} &= C'_{it}\left(D_{it}\left(p^*_{it}, p^*_{-it}\right)\right) - \frac{1}{\frac{\partial D_{it}\left(p^*_{it}, p^*_{-it}\right)}{\partial p_{it}} \frac{1}{D_{it}\left(p^*_{it}, p^*_{-it}\right)}}. \end{aligned}$$

The empirical specification *(2.1)* can be derived from *(2.7)* by assuming that

(1) Equilibrium prices are observed with an $i.i.d.$ error: $p_{it} = p^*_{it} + \nu_{it}$,

(2) For each station the marginal cost is proportional to the observed wholesale price and does not depend on the equilibrium quantity: $C'_{it}\left(D_{it}\left(p^*_{it}, p^*_{-it}\right)\right) = \theta_i c_t$,

(3) The semi elasticity of station i's demand does not depend on the equilibrium price vector and is a station specific constant: $-\frac{1}{\frac{\partial D_{it}\left(p^*_{it}, p^*_{-it}\right)}{\partial p_{it}} \frac{1}{D_{it}\left(p^*_{it}, p^*_{-it}\right)}} = \alpha_i$.

These assumptions imply that unlike the differentiated Bertrand oligopoly, the price equilibrium implied by *(2.1)* is not a strategic one. Instead, it captures the behavior of a monopolist with constant semi-elasticity of demand and constant marginal cost in time t. The most general specification and therefore its restricted versions can simply be interpreted as reflecting monopoly pricing at different levels of heterogeneity: stations, groups of stations defined by observables and market level. Nevertheless, empirical and theoretical papers typically describe the gasoline retail market by some form of Bertrand competition with varying degree of local market power. Verlinda (2008) for example suggests that interacting station characteristics with cost and price changes in the ECM captures the effect of market power on pricing asymmetry. Such an interpretation is only valid if one is willing to assume restrictions on the parameters that separate station heterogeneity from market specific variables that aim to capture the strategic effects of competition.

To do this, however, one needs to put some functional form on *(2.7)* by assuming an explicit form for the demand and the marginal cost function. An example of this could be the Bertrand–Shubik demand system in which equilibrium prices for each station i in time t can be expressed as (see for example Wang and Zhao, 2007):

$$p_{it}^{*} = \theta_1 \left(n_{it}^{m}, \gamma_{it}^{m}\right) V_{it}^{m} + \theta_2 \left(n_{it}^{m}, \gamma_{it}^{m}\right) \overline{c}_{it}^{m} + \theta_3 \left(n_{it}^{m}, \gamma_{it}^{m}\right) c_{it}. \tag{2.8}$$

where θ_1,θ_2,θ_3 are non-linear functions of n_{it}, the number of i's competitors and γ_{it}, a demand parameter. V_{it} is the total demand in station i's local market, $\overline{c}_{it}$ is the average marginal cost in station i's local market and c_{it} is station i's marginal cost. n_{it}^{m}, γ_{it}^{m}, V_{it}^{m} and $\overline{c}_{it}^{m}$ are all market specific variables, therefore they are the same for all stations in market m. These variables could be used to identify the competitive effects in this model, because station specific heterogeneity is confined to the station-specific marginal cost c_{it}. Unfortunately, γ_{it} and V_{it} are unobserved and in practice only a uniform wholesale price is observed: $c_t = \overline{c}_{it} = c_{it}$. If one assumes that station i's marginal cost can be expressed as $c_{it} = \delta_i c_t$, then the problem of identifying competitive effects boils down to how to separate δ_i (station heterogeneity) and $\theta_3 \left(n_{it}^{m}, \gamma_{it}^{m}\right)$ (market structure parameters).

The three specifications outlined previously offer different solutions to the identification problem. The fully heterogenous estimation uses the two step mean group estimator. In the first step, station by station regressions are estimated. In the second step, these parameter estimates are regressed on market and station specific categorical variables to provide mean estimates for the groups defined by these variables. The observed heterogeneity specification makes much stronger restrictions by assuming that it is possible to separate individual heterogeneity and competitive effects into two sets of linear parameters: the first related strictly to firm specific observables (brand) and the second related to market specific observables (number of firms, type of competitors). The first set is interpreted to capture individual heterogeneity, while the second set is interpreted as indicating competitive effects. Finally, the homogenous panel specification simply cannot address this issue since it does not allow for any heterogeneity in the wholesale price parameter. However, it serves as a useful benchmark that can be compared easily with classic results based on time series data.

There is no unique way to separate individual heterogeneity from market level competitive effects. Verlinda (2008) estimates a random parameter version of *(2.2),* in which mean parameters are shifted by market level observables (for example the number of firms). These mean effects are interpreted to capture effects of competition (market power). Comparing this paper's empirical strategy to Verlinda's (2008) method boils down to a trade off between a semiparametric two stage estimation (mean group estimator) and a parametric joint estimation (hierarchical-Bayes estimator) that relies on a normality assumption for the random parameters.

2.5. Estimation and results

Results are presented in four steps. First, the equilibrium pricing relationship estimates are introduced briefly. Second, the aggregate results are discussed to confirm whether the hypothesis of no asymmetry holds at the market level. Third, the heterogeneity of the retail price response in the week of the wholesale price change is studied. This instantaneous response is likely to be the most relevant for consumers, because the bulk of the adjustment takes place within this first week. Finally, the heterogeneity in the full dynamic response is analyzed by comparing cumulative response functions following BCG (1997).

2.5.1. The equilibrium pricing relationship

The equilibrium pricing relationship describes the static price setting behavior of stations and it is estimated as a cointegrating relationship. Table 4 presents the average estimated coefficients for the three specifications. These are calculated as the average marginal effects of the wholesale price for the observed heterogeneity panel estimates and as the unweighted average of the estimated parameters for the station by station regressions. This latter is also referred to in the literature as the mean group estimator (Pesaran and Smith, 1995).[4] All three specifications are estimated by dynamic OLS using one lead and two lags in order to control for autocorrelation in the disturbance. Moreover, the panel estimators also include fixed effects.

Table 4: Aggregate price equilibrium parameters

	Fully homogenous panel	Observed heterogneity	Unobserved heterogeneity (1)
Wholesale price coefficient	1.02**	1.01**	1.02**
Number of stations	1291	1291	1291
Number of observations	53685	53685	53685
R2 (within)	0.99	0.99	0.99
Estimator	Fixed effects, dynamic OLS	Fixed effects, dynamic OLS	Dynamic OLS

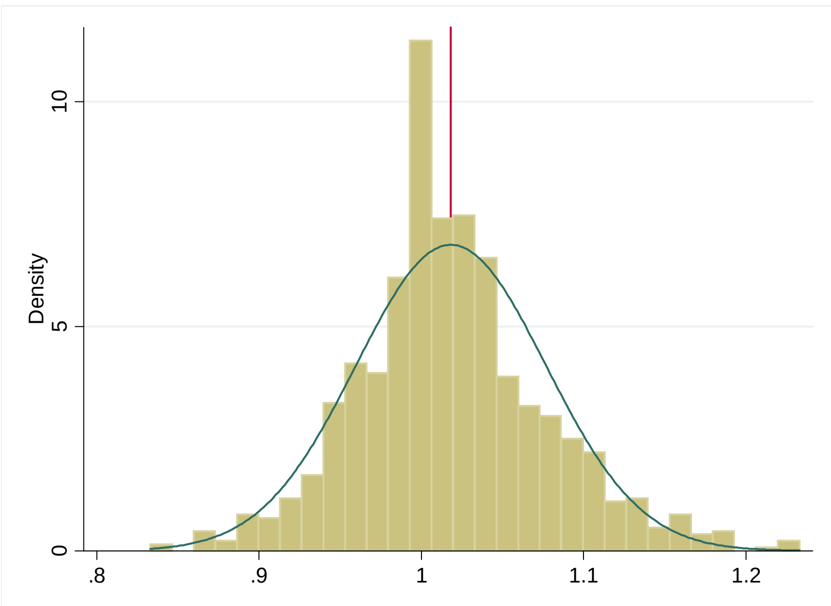
(1) Mean group estimator.

[4]Standard errors are clustered by stations for the homogenous panel and observed heterogenity estimations. The variance of the mean group estimator $(\widehat{\theta}^{mg})$ is caclulated as $var(\widehat{\theta}^{mg}) = \frac{1}{N(N-1)}\sum_{i=1}^{N}\left(\widehat{\theta}_i - \widehat{\theta}^{mg}\right)^2$.

The mean estimates are practically the same in all three specifications and indicate full long term pass-through. They are also consistent with evidence from aggregate data (for example BCG,1997) that other station-level studies (Lewis, 2004 or Verlinda, 2008) could not replicate. The mean coefficient of one is consistent both with a pricing strategy that sets a constant mark-up over the wholesale price.

Although the mean coefficients in all three specifications are the same, there is sizeable heterogeneity at lower levels of aggregation. This dispersion is demonstrated best with the station by station estimates of the unobserved heterogeneity specification. Graph 2 shows the parameter distribution of this estimation, which is remarkably close to normal, except for the concentration of estimates at unity. Most estimates range from 0.8 to 1.2. This variation can reflect the unobserved heterogeneity both in the true wholesale costs of stations and in the residual demand they face.

Graph 2: Wholesale price parameter distribution - unobserved heterogeneity estimates



2.5.2. Aggregate asymmetry

The error correction model is estimated based on the Engel-Granger two step procedure, where the error correction term is the residual of the price equilibrium regression. The descriptive statistics and the public information about the wholesale prices both suggested that there should be no sizeable asymmetry in retail price responses. Table 5 shows the average estimates of asymmetry parameters from the three ECM specifications. Similarly to the price equilibrium regressions, average coefficients are the average marginal effects in the sample for the panel estimates and the mean group estimates for the unobserved heterogeneity specification.[5] Estimates of the average asymmetry parameters are remarkably close to each other in all three specifications.

Table 5: Aggregate price asymmetry estimates

	Fully homogenous panel	Observed heterogneity	Unobserved heterogeneity (1)
Positive wholesale change	0.99**	0.99**	0.98**
lag1	0.016**	0.022**	0.011
lag2	0.007**	0.009**	0.005
Negative wholesale change	0.92**	0.93**	0.92**
lag1	-0.009**	-0.010**	-0.015
lag2	-0.013**	-0.008**	-0.008
ECM term	-0.17**	-0.20**	-0.27
Number of stations	1,289	1289	1287
Number of observations	51719	51719	51717
R2 (within)	0.87	0.88	0.93

(1) Mean group estimator.

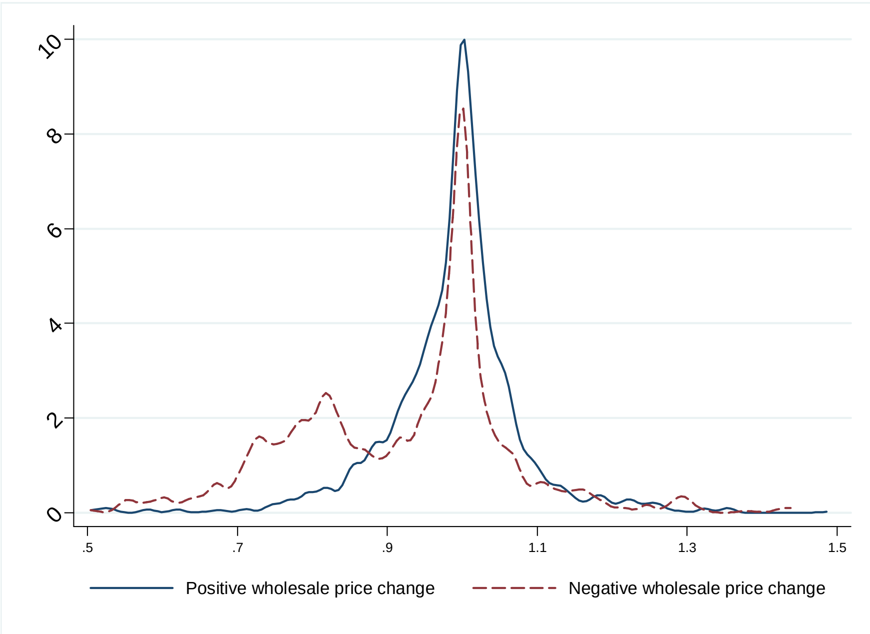
The results from the ECM-s imply that on average there is a statistically significant but economically unimportant asymmetry among retail price responses to positive and negative wholesale price shocks. While the retail price adjusts virtually instantaneously to a positive wholesale price change, negative shocks are not

[5]Standard errors are the same as in the equilibrium price regressions.

passed on fully in the same week. Moreover, while for positive changes the lagged responses suggest some slight overshooting, for negative changes they indicate slow adjustment.

The speed of adjustment parameter is of similar size as the largest lagged effects of wholesale price changes. They imply that the half life of a random shock is 2 to 4 weeks. The unobserved heterogeneity average estimate implies the fastest convergence to equilibrium, while the homogenous panel estimate implies the slowest. The speed of adjustment parameter is not significant for the unobserved heterogeneity specification, showing that the mean group estimator is less efficient.

Graph 3: Distribution of the instantaneous asymmetry parameter



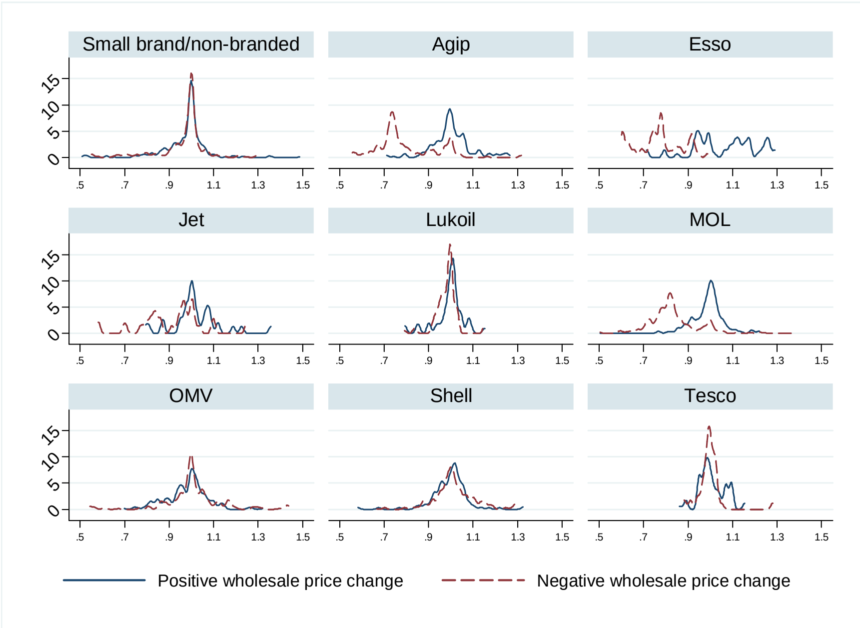
There is nevertheless substantial heterogeneity behind these aggregate results. Graph 3 draws the histograms of positive and negative instantaneous response parameters, from the unobserved heterogeneity specification. The two distributions are significantly different, based on a Kolmogorov-Smirnov test one can reject the null hypothesis of equal distributions with probability 1. Moreover, the graph replicates the insight for the retail-wholesale price change ratios: the parameter distribution of the negative response is bimodal. There are stations (centered around 1) that do not show any sign of asymmetry, there are however others (centered around 0.8) that show significant asymmetry. The aggregate result for the negative transmission parameter is basically the average of these two groups. The next section tries to separate these two groups based on the observable characteristics of stations.

2.5.3. Instantaneous response parameters: brand matters

Since most of the retail price adjustment takes place within the week of the wholesale price change, this section focuses on the heterogeneity in the instantaneous response parameters. First, the distribution of negative and positive wholesale price parameters is shown for different characteristics using the unobserved heterogeneity estimates. The aim is to see whether there is a characteristic that separates clearly the stations around the two modes of the negative wholesale price change parameter distribution. Graph 4 shows the parameter distributions by brand. For small brands and non-branded stations, Lukoil, OMV and Shell, there is no difference between positive and negative parameter distributions and there is no bimodality. For Agip, Esso, and MOL there is a clear mean shift between the negative and the positive parameter distributions and there is a slight sign of some bimodality. Jet and Tesco are a less clear-cut case.

The instantaneous parameter distributions conditioned on the number of competitors (Graph 5) or on counties (Graph 6) does not show a similarly clear pattern. Most of these conditional distributions show the same bimodality as the unconditional distributions in Graph 3 and none of them demonstrates a clear mean shift between positive and negative parameter distributions.

Graph 4: Distribution of the instantaneous asymmetry parameter by brand



Graph 5: Distribution of the instantaneous asymmetry parameter by number of competitors

Graph 6: Distribution of the instantaneous asymmetry parameter by county

Looking at these conditional parameter distributions, however, gives only a crude assessment about the relative importance of these different characteristics. Therefore, Table 6 presents the partial effects of observed characteristics for both the observed and the unobserved heterogeneity specifications. These partial effects show how the characteristics modify the instantaneous price response to a 1 HUF wholesale price change. The comparison group is a non branded station in county 1 without any competitors. This base group does not show any sign of asymmetry. In order to compare these estimates to the aggregate results, the first line of the table repeats the latter for the asymmetry in the instantaneous response. The first two columns for each specification show the partial effects of the positive and the negative wholesale price changes. The third column shows the differences of these parameters as an indicator for the size of the asymmetry. For the observed heterogeneity specification the parameters are the interactions between the instantaneous wholesale price change and the characteristics. For the unobserved heterogeneity specification these parameters are the results of the second stage regression. Standard errors are clustered by stations for the observed heterogeneity specifications and are heteroscedasticity robust standard errors for the unobserved heterogeneity specification.

Table 6. Partial effects of observed characteristics

	Positive wholesale price change	Negative wholesale price change	Asymmetry	Positive wholesale price change	Negative wholesale price change	Asymmetry
Overall average	0.985	0.928	0.06	0.984**	0.920**	0.06
Brand						
Agip	0.04**	-0.10*	0.14	0.06**	-0.13**	0.18
Esso	0.10**	-0.15*	0.25	0.11**	-0.18**	0.29
Jet	0.08**	-0.06*	0.13	0.05**	-0.04	0.09
Lukoil	0.05**	0.03	0.02	0.05**	0.04**	0.01
MOL	0.04**	-0.11*	0.14	0.04**	-0.12**	0.16
OMV	0.04**	0.08*	-0.04	0.04**	0.08**	-0.04
Shell	0.05**	0.08*	-0.03	0.06**	0.08**	-0.02
Tesco	0.05**	0.03	0.02	0.05**	0.05**	0.00
Number of competitors						
1	0.03	0.01	0.02	0.02	0.00	0.03
2	0.02	-0.02	0.04	0.00	0.00	0.01
3	0.01	-0.01	0.02	0.01	0.00	0.01
4	0.02	-0.01	0.03	0.01	0.00	0.01
5	0.02	0.00	0.02	0.00	-0.01	0.01
6	0.01	0.00	0.01	-0.01	0.01	-0.02
7	0.01	0.02	-0.01	0.01	0.00	0.01
8	0.03	-0.05	0.08	0.02	-0.04	0.06
9	0.01	0.03	-0.03	0.01	0.00	0.01
10	0.00	-0.05	0.04	-0.02	-0.03	0.01
Type of competitor						
Small brand / non-branded	0.00	0.00	0.00	0.00	0.00	0.00
Agip	0.00	0.03**	-0.03	0.01	0.02**	-0.02
Esso	0.00	0.04**	-0.04	0.00	0.00	0.00
Jet	0.00	-0.01	0.01	-0.02**	0.00	-0.02
Lukoil	0.00	-0.01	0.01	-0.01	-0.01	0.00
MOL	-0.01	0.01	-0.02	0.01	-0.01	0.02
OMV	0.01	0.00	0.01	0.02**	0.01	0.01
Shell	0.01	0.00	0.00	0.02**	0.00	0.02
Tesco	0.00	-0.02	0.02	0.01	-0.01	0.02
County						
2	0.03	-0.01	0.03	0.00	-0.03	0.03
3	0.00	0.04	-0.04	-0.02	0.02	-0.05
4	-0.02	0.00	-0.02	-0.02	-0.01	-0.02
5	0.03	-0.05	0.08	0.01	-0.02	0.04
6	-0.01	-0.01	-0.01	-0.04*	-0.02	-0.02
7	-0.01	0.00	-0.01	-0.03*	-0.03**	0.00
8	0.03	0.01	0.02	-0.01	-0.01	0.00
9	0.02	0.03	-0.01	0.01	0.02	-0.01
10	-0.04**	0.06**	-0.10	-0.03*	0.05**	-0.08
11	0.00	0.03	-0.04	-0.02	0.01	-0.04
12	0.00	0.00	0.01	-0.01	-0.03**	0.02
13	-0.02	0.08*	-0.10	-0.04**	0.04	-0.08
14	0.01	0.00	0.01	-0.01	0.00	0.00
15	-0.02	0.00	-0.02	-0.03*	0.01	-0.03
16	-0.02	0.01	-0.04	-0.03**	0.00	-0.03
17	0.03	0.01	0.03	0.00	-0.03	0.03
18	-0.02	0.00	-0.02	-0.04**	0.00	-0.04
19	0.00	0.02	-0.02	-0.01	0.01	-0.02
20	-0.02	0.01	-0.03	-0.02	0.00	-0.02
Number of stations		1289			1289	
Number of observations		51719			51717	

In line with the evidence from the conditional distributions, some brands (brand Agip, Esso, Jet and MOL) show much higher asymmetry in the instantaneous response than the aggregate. They indicate that in response to a 1 HUF wholesale price decrease, the instantaneous retail price reductions are 13 to 25% (9 to 29%) smaller than the corresponding increases. These are magnitudes that are of economically significant size. Taking the average wholesale price change of 4 HUF the largest asymmetry (25%) implies that stations of brand 3 decrease prices only with 3 HUF while they increase prices with 4 HUF in the first week. This 1 HUF difference is approximately 7% of the markup, which is arguably a significant amount for the firm. Compared to the average prices, however, even these asymmetries are irrelevant. The importance of the brand of the stations in explaining the variation in asymmetry is not straightforward to interpret. The brand of a station can capture many unobserved features of a station and therefore it is an ideal control for heterogeneity.

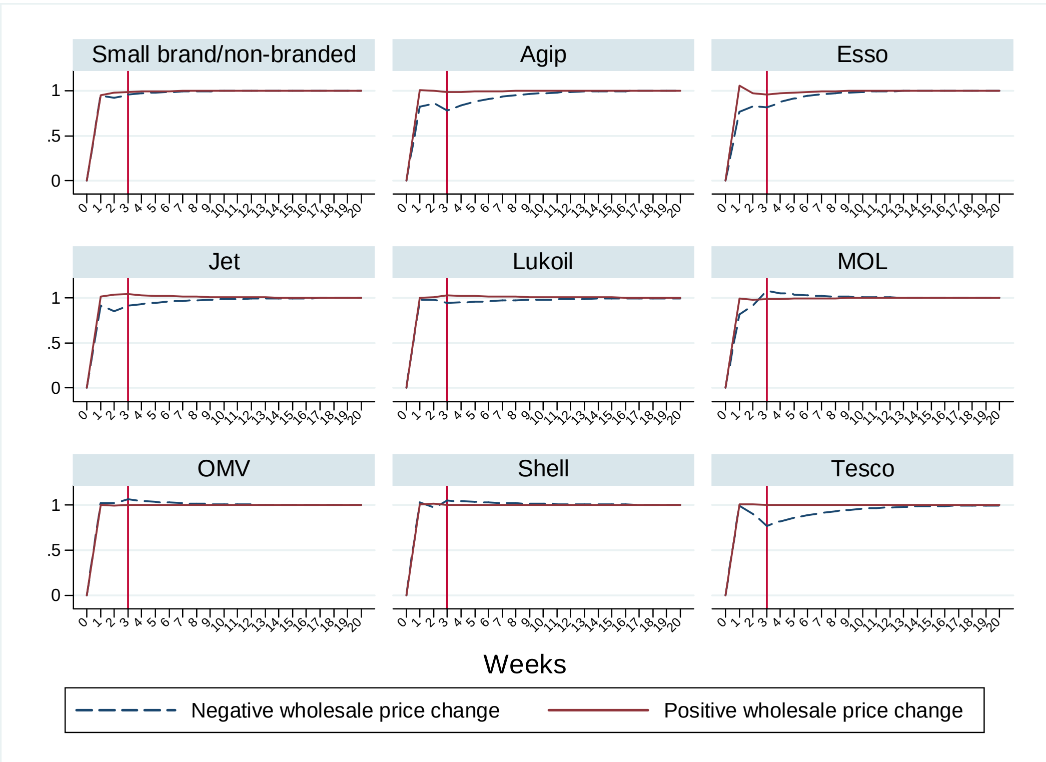
While partial effects of brands can be estimated precisely, most parameters of the other characteristics are not significantly different from zero. The partial effects are small for the competition related variables (number of competitors and type of competitor). The effect of geographic location, which is mostly capturing demand heterogeneity, is on average the same size. However, two counties (10 and 13) have relatively large dampening effect on asymmetric behavior.

2.5.4. Dynamic adjustment

While the instantaneous response captures the largest part of the retail price adjustment, the full extent of asymmetry can only be judged based on the dynamic response to a unit change in the wholesale price. The previous section indicated that it is the brand of the stations that influences this adjustment the most, therefore the dynamics are presented separately for each brand.

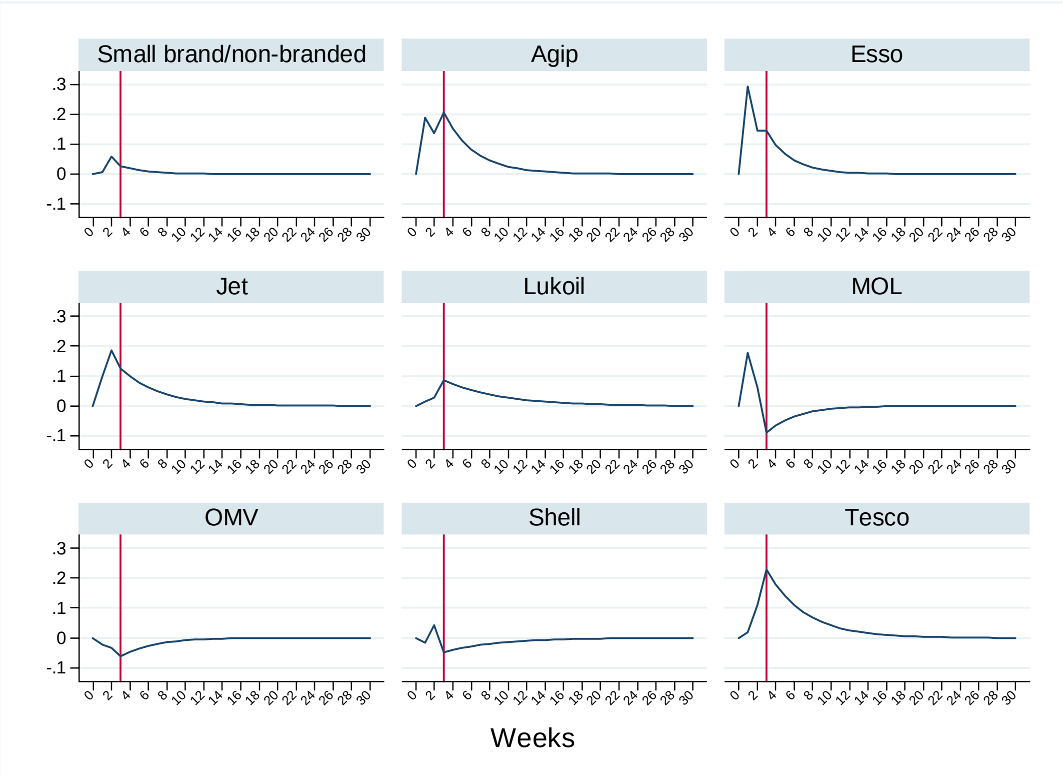
The adjustment path is shown in Graph 7 that presents the cumulative response functions (CRF). The CRFs show the evolution of the retail price in response to a 1 HUF wholesale price increase/decrease. To ease comparison, for negative price changes the absolute values are depicted. The CRFs are based on the partial effects estimated from the second stage of the unobserved heterogeneity specification (as shown in the Appendix, results are practically the same when using the observed heterogeneity estimates). Therefore, the differences among the CRF are due to brand differences while holding other observable characteristics unchanged. The comparison group is the average unbranded station, which is presented in the first panel.

Graph 7: Cumulative response functions - unobserved heterogeneity



The dominant feature of the CRFs is the quick adjustment in the first week and the convergence to the long term equilibrium, which is 1 in this case. There are noticeable differences between the adjustment paths for positive and negative wholesale changes for Agip, Esso, Jet and Tesco. MOL is the only one that shows sizeable overshooting for negative wholesale changes. Complete convergence takes place in about 16 weeks for most brands. It is useful to keep in mind though that on average there was a price change every third week in the sample (vertical line), therefore it is unlikely that the full adjustment path is completed.

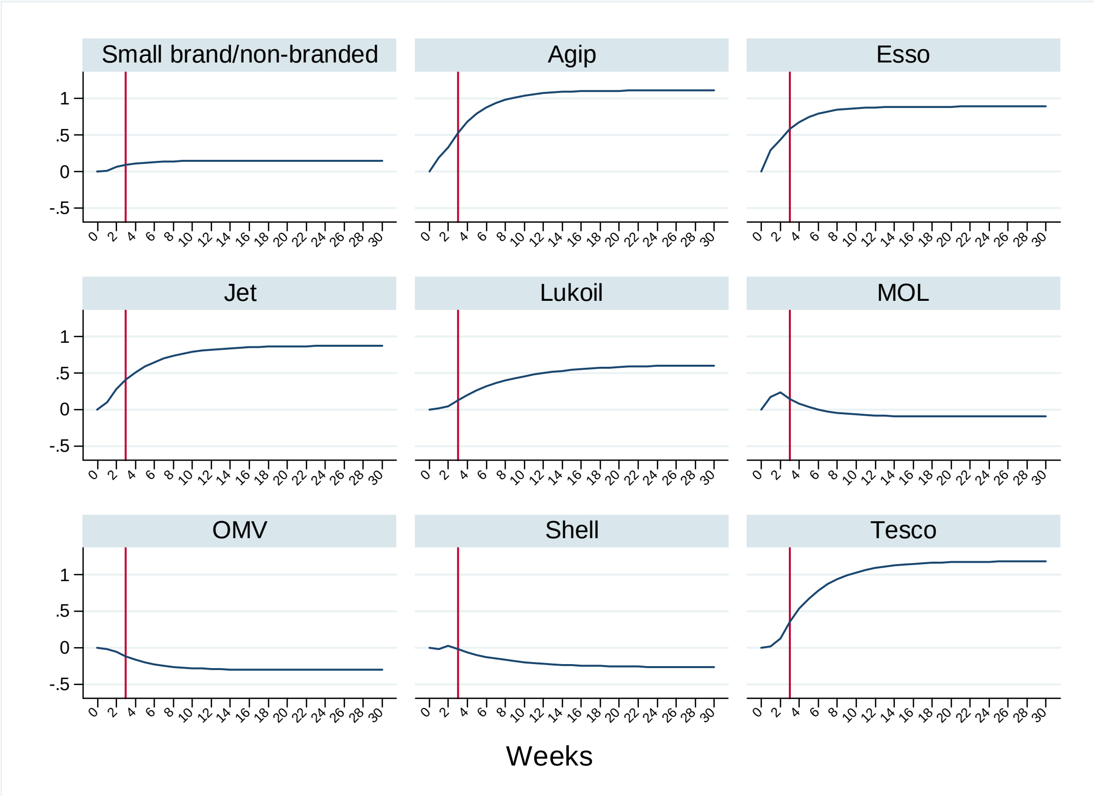
Graph 8 Difference between negative and positive response functions - unobserved heterogeneity



While the CRF graphs are informative about the dynamic adjustment in general, it is not straightforward to read the size of the asymmetry from them. Therefore Graph 8 shows the difference between the positive and the negative CRFs. Agip, Esso, Jet and Tesco show sizeable asymmetry in the first weeks following the wholesale price change. Asymmetry values for all these brands reach 0.2 and stay above 0.1 within the first three weeks. Agip and Tesco reach the peak in week 3 while Esso and Jet are reducing their asymmetry after the first and second week respectively. Tesco did not show sizeable instantaneous asymmetry but the CRF reveals that in case of negative wholesale shocks there is a reversal towards higher prices between the first and the third week. Conversely, MOL is characterized by relatively large instantaneous asymmetry, but Graph 8 shows that there is a strong reduction in this asymmetry that even reaches negative values for the third week.

The CRF differences clearly show that for four brands asymmetry is sizeable in the first weeks following a wholesale price change. To see what the hypothetical economic cost of these asymmetries is, Graph 9 presents the integrals of the CRF differences. These describe for each week the cumulative difference between the CRFs and express how much more a consumer paid for a unit price increase than what he saved on a unit price decrease (assuming he buys gasoline each week). Or from the firm's perspective, how much more it earns on a unit price increase compared to a unit price decrease. For brands Agip, Esso, Jet and Tesco this cumulative difference converges to approximately 1 HUF. However, this level is unlikely to be reached, since on average the wholesale price changes faster than convergence could play out. After three weeks, which is the average duration between wholesale price changes, the cumulative difference is approximately 0.5 HUF for these brands. This magnitude means that compared to the instantaneous response, the transmission asymmetry doubles to 14% of the markup during the first weeks of dynamic adjustment. Asymmetry is, nevertheless, still tiny compared to consumer prices.

Graph 9: Integral of the difference between negative and positive response functions - unobserved heterogeneity



2.5.5. Brands that price asymmetrically

The results indicated that it is the Agip, Esso, Jet and Tesco stations that adjust retail prices asymmetrically. Looking back at Tables 1 and 2 in the data description part, one can see that these brands had market shares below 10% in 2007. Their average margins, however, differ quite substantially: Agip is a high margin brand with 5.7%, while Esso and Jet are among the lowest margin brands. If one is willing to assume that margins indicate demand elasticities, then this latter finding means that these brands face quite different residual demands. Therefore, demand heterogeneity is unlikely to explain why this group of brands prices asymmetrically. These brands also differ in how often they are competitors of other stations: while

64% of the stations have an Agip competitor, only 35% and 38% have Esso or Jet competitors respectively. This suggests that also the competitive position of these brands is quite different. Tesco's characteristics are between these two extremes.

One possible explanation for the similarity in pricing behavior would be that these brands are competitors of each other. Table 7 shows the percentage of stations within a brand (rows) that have a given brand as a competitor (columns). The last row shows the average percentage of stations that have an other given brand as competitor. The Esso (row) - Agip (column) cell indicates that 88% of Esso stations have an Agip station as competitor. Comparing this to the last figure in the same column (61%) reveals that compared to the average Esso stations are more likely to have an Agip competitor. The same is true for Jet stations but not for Tesco stations. For Agip all other asymmetrically pricing brands are relatively infrequent competitors, while Jet and Esso are each others' most frequent competitor. Generally if one looks at the differences of competitor frequency by brand from the overall average frequency, it turns out that Agip and Tesco are located in less populated markets and therefore have on average less competitors from each brand than the average, while the opposite is true for Esso and Jet.

Table 7. Percentage of stations within a brand that have a given brand as a competitor

Brand of station / brand of competitor	Non-branded	Agip	Esso	Jet	Lukoil	MOL	OMV	Shell	Tesco
Non-branded	.	49	27	29	33	91	60	67	32
Agip	76	.	34	37	42	94	60	69	39
Esso	98	88	.	73	69	100	96	96	59
Jet	97	83	59	.	55	97	92	98	60
Lukoil	77	56	21	26	.	92	57	73	38
MOL	80	57	31	33	39	.	60	68	40
OMV	93	75	55	56	58	97	.	86	59
Shell	89	71	49	51	54	96	75	.	60
Tesco	82	56	23	24	29	97	58	80	.
All stations	84	61	36	39	43	95	64	73	45

2.6. Conclusion

The paper estimated an ECM model using gasoline station level panel data from Hungary to provide additional evidence on asymmetric pricing practices in gasoline retail markets. The main results are the following. Wholesale price changes are fully passed through to retail prices in the long run, since the wholesale price parameter estimate in the pricing equation is 1. This means that in the long run, mark-ups are independent of wholesale prices in this market. This finding is consistent with constant mark-up pricing or monopoly pricing as well.

When looking at the dynamic adjustment of retail prices to wholesale price changes one finds that on aggregate there is no pricing asymmetry. Therefore, retail prices practically increase at the same speed in response to a wholesale price increase as they decrease when the wholesale price is decreasing. Moreover, on average, almost all of the adjustment takes place in the first week. The finding of no pricing asymmetry on aggregate confirms the null hypothesis that was based on the publicly available wholesale price information.

The aggregate result of no asymmetry, however, masks significant heterogeneity at station level. A sizeable fraction of stations does price asymmetrically and these stations earn on average 0.5 HUF more on every unit of wholesale price increase compared to a reduction. This asymmetric pricing is explained by the brand of the station, which is the level of price setting in this market. Brands that price asymmetrically have small market shares and are not vertically integrated. Brands with large market share, vertically integrated brands and non-branded stations do not price asymmetrically. Other, local market specific variables like number of stations or type of competitors do not explain asymmetric pricing.

Which theoretical explanations are supported by these results? Table 8 gives an overview of this question. The discussion on identification suggested that the brand of a station is best thought of as a proxy that captures unobserved station heterogeneity. It can proxy both unobserved marginal cost differences of firms and unobserved taste heterogeneity that can influence marginal costs through the equilibrium quantity demanded. The fact that this variable is key in explaining pricing asymmetries underlines the importance of controlling for unobserved heterogeneity in studies using station level data. Furthermore, the finding that the asymmetrically pricing brands are small, are not vertically integrated and their average margins differ considerably suggests that the explanation for the asymmetric pricing could be related to the cost structure of these brands. Small brands with low station density can have higher transportation costs and inventory adjustment cost than larger brands because of the indivisibility in transportation loads.Nevertheless, verifying this hypothesis would require more detailed data on the cost structure of gasoline stations and information about quantities.

Tacit collusion and search-based explanations are not supported by the results. The significant asymmetry appears to be not at market level, but at brand level: in the same market there are brands that price asymmetrically and brands that do not. Therefore tacit collusion type explanations that focus on market interaction are rejected by the results. Also the finding that market structure variables do not explain asymmetries and that the brands with the largest market shares set prices symmetrically provides additional evidence against tacit collusion type explanations. The search friction explanation would imply aggregate retail price asymmetry since search frictions are not confined to specific local markets. However, the results do not support this prediction. Moreover, search frictions are unlikely to be brand-specific either, therefore search based models are also not supported by the result that asymmetric retail price adjustment is confined to specific brands.

Table 8. Theories and findings

Theory/Result	Tacit collusion	Inventory/ adjustment costs	Search frictions
No aggregate asymmetry	0	+	-
Small brands price asymmetrically	-	+	-
Large brands price symmetrically	-	+	-
Number and type of competitors irrelevant	-	0	0

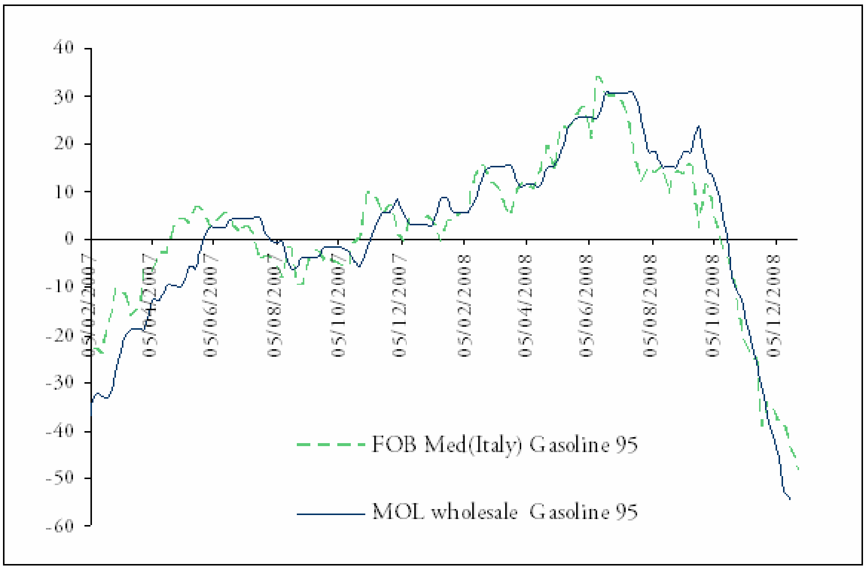
+: consistent with the theory, -: does not support the theory, 0: no prediction from theory,

To sum up, the results draw attention to the heterogeneity in asymmetric pricing behavior across brands. This finding rejects collusion and search model based explanations, but is consistent with explanations related to the gasoline retail firm's adjustment costs (transportation and inventories). The importance of the firm's costs different from the wholesale price suggests that future research should focus on collecting more information about these costs in order to provide more robust evidence on asymmetric retail price responses.

2.7. Appendix

Graph 10: List wholesale price of MOL and the market wholesale price, HUF/l

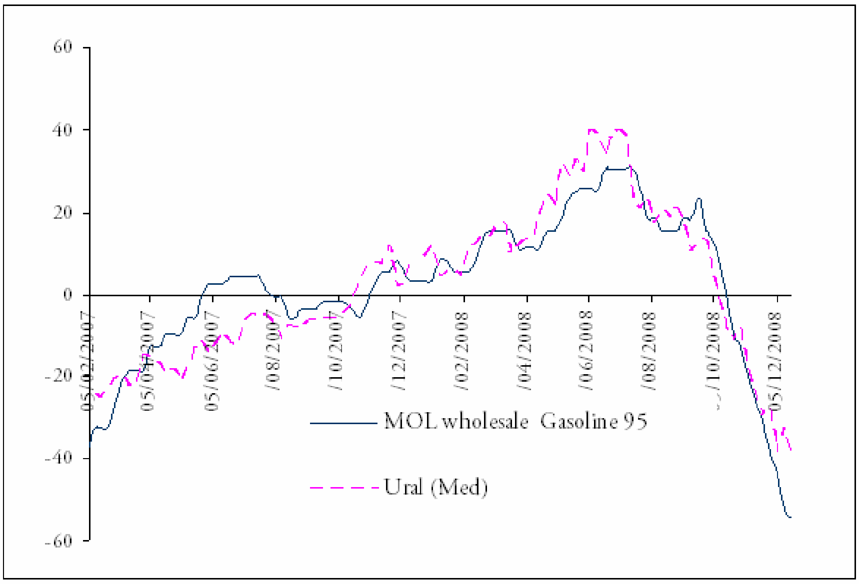
(deviations from the respective averages)



Source of market price data: Energy Center,
http://www.energiakozpont.hu/download.php?path=files/energiastatisztika/Koolajpiaci_informaciok.xls

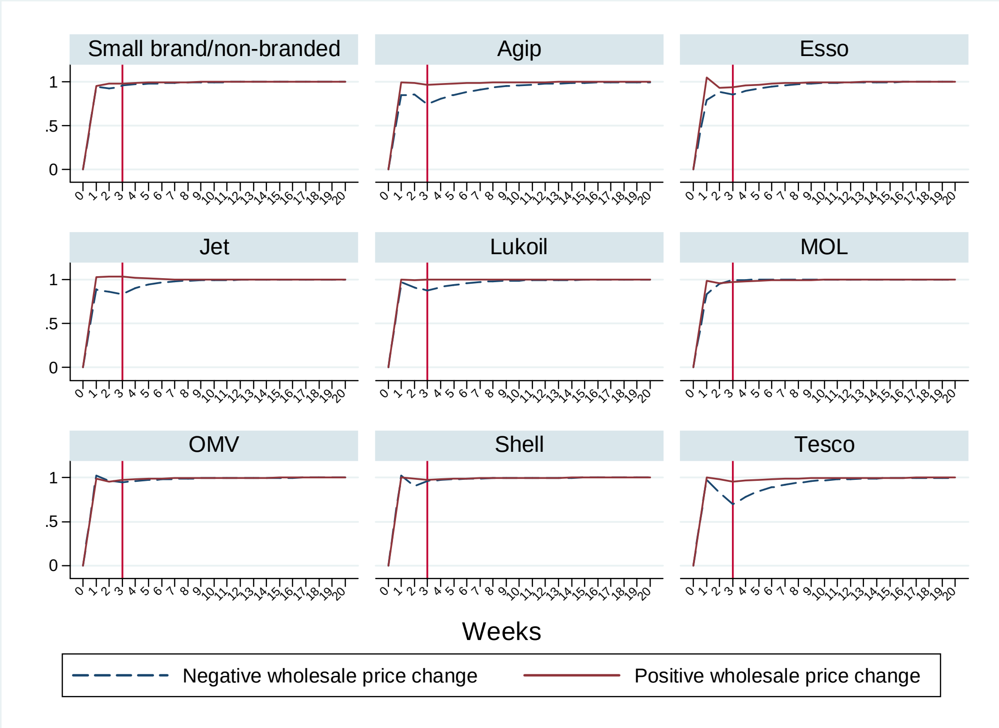
Graph 11: List wholesale price of MOL and the oil price (Ural), HUF/l

(deviations from the respective averages)

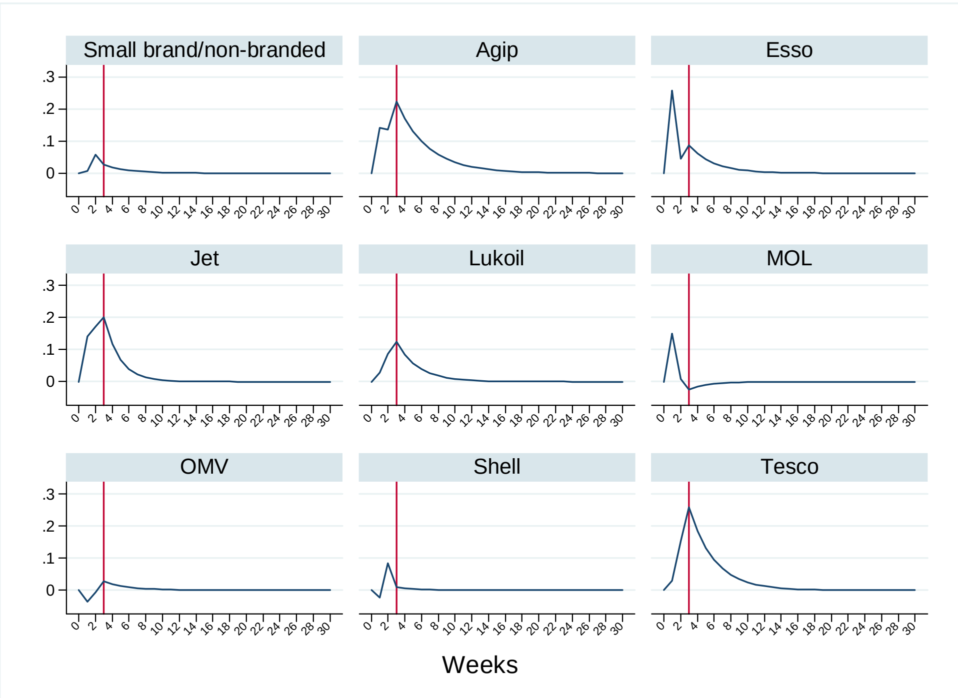


Source of oil price data: Energy Center,
http://www.energiakozpont.hu/download.php?path=files/energiastatisztika/Koolajpiaci_informaciok.xls

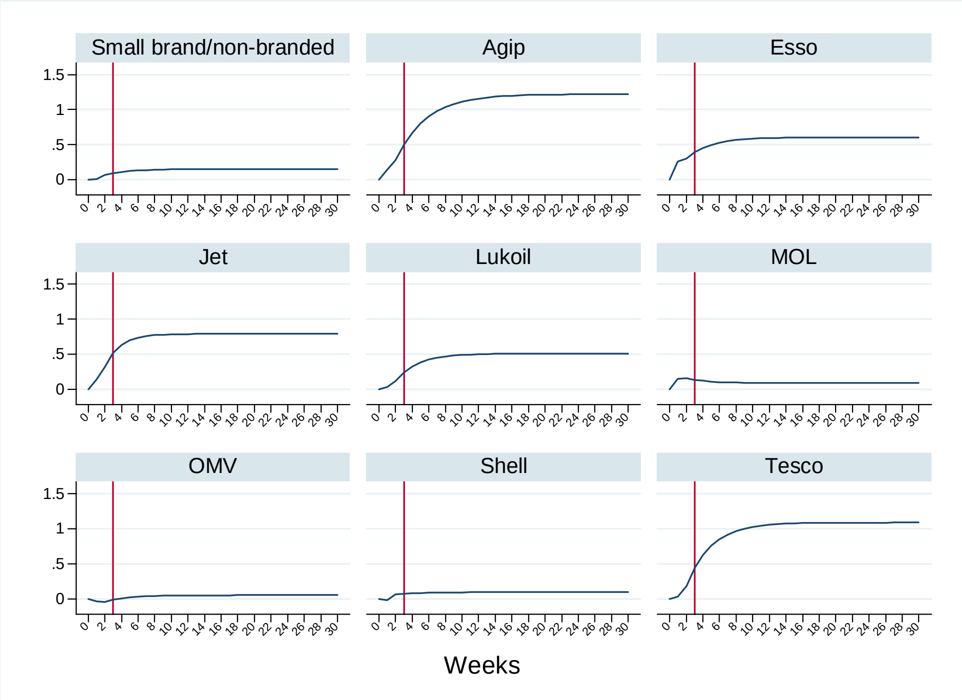
Graph 12: Cumulative response functions - observed heterogeneity



Graph 13: Difference of negative and positive response functions - observed heterogeneity



Graph 13: Integral of the difference between negative and positive response functions - observed heterogeneity



CHAPTER 3

Separating the ex post effects of mergers: an analysis of structural changes on the Hungarian retail gasoline market

Joint with Gergely Csorba[1] and Dávid Farkas[2]

3.1. Introduction

From the late 1990s, there has been a growing need to evaluate the performance of antitrust policies, and especially whether mergers contributed to the observed price increases in their respective industries.[3] The goal of ex post merger evaluation is to identify the price change due to the merger itself and separate it from the price effect of any other economic factors such as changes in demand and cost conditions. The central question of most previous studies was to find the total (average) price effect attributed to the merger, but less attention has been given to analyzing the difference in the effects a merger can have on the various firms affected by the mergers. In this paper, we apply difference-in-differences methods to identify the price effects of simultaneous mergers, and break down the total effect of each merger by separating the effects on the prices of the buyer and seller firms and on the prices of their respective competitors.

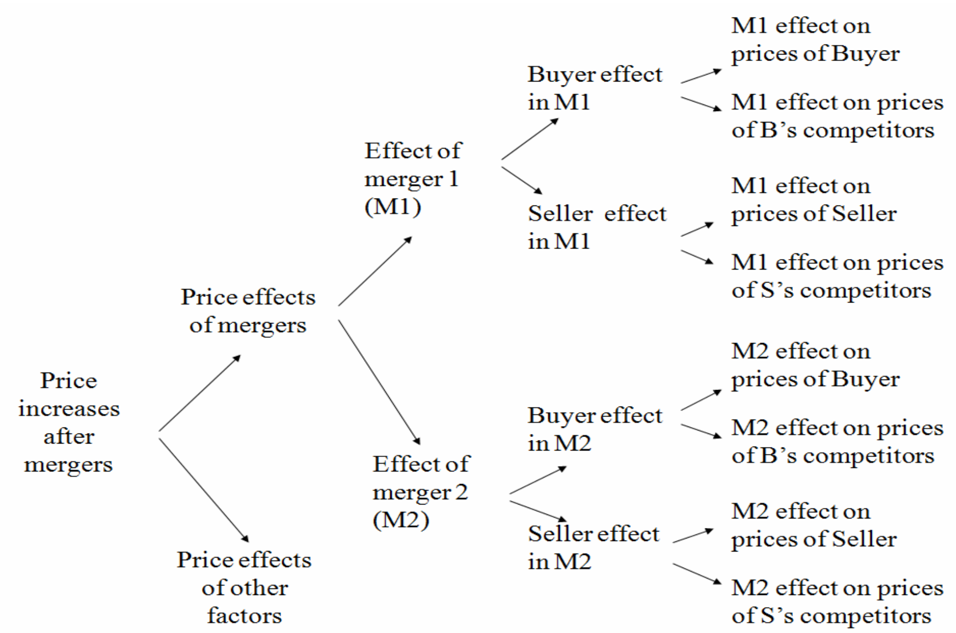
[1]Hungarian Competition Authority (GVH) and Institute of Economics (Hungarian Academy of Sciences)

[2]Hungarian Competition Authority (GVH)

[3]See the LEAR (2006) report prepared for the DG Competition European Commission, Office of Fair Trading and Competition Commission (2005) and contributions to the "Measuring the Economic Effects of Competition Law Enforcement" conference organized by the Dutch Competition Authority (NMa) in 2007, which were published in the December 2008 issue of De Economist.

Graph 1 illustrates the eight different merger effects we aim to separate in our ex post evaluation of two mergers.

Graph 1. The different merger effects to separate



Separating these effects enables us the testing of some important predictions of academic and antitrust literature, which argue that a merger can result in different price changes for different firms, depending on their role in the merger. First, the most robust prediction is that a merger will result in a larger change in merging firms' pricing than in competitor firms' pricing as the former can fully internalize the effect of eliminating the competitive constraint (externality) the two firms had on each other before the merger.[4] Second, in mergers with local markets, a larger price increase is expected on markets where both merging firms are present (or are closer competitors to each other), since the merger removes a direct competitive constraint between their respective outlets.[5] Third, a merger might have a different

[4]See the classical Davidson and Deneckere (1985) price competition model with differentiated products or Vives (1999) on more general results in various oligopolistic settings.
[5]Levy and Reitzes (1992) develop a merger model of spatially differentiated firms leading to this result.

effect on the two firms involved, as the business policies and supply conditions of the firms will likely converge towards each other, and the change is usually conjectured to be larger for the case of the acquired firm than for the buyer firm.[6]

In this paper, we provide an ex post assessment of two mergers on the Hungarian retail gasoline markets, which happened almost simultaneously: the acquisition of Jet by Lukoil in February 2007 and the acquisition of Esso by Agip in July 2007.[7] Our detailed panel on station-level prices offers an intuitive way to estimate the price effects of the two mergers with difference-in-differences methods, where we exploit the variation in the presence of merging firms across local markets. Because we observe almost all possible combinations of the four firms' stations in distinct local markets, we can form different treatment-control group pairs to identify separate effects for each merging firm and their competitors.[8]

As a preliminary illustration of our results, Table 1 shows the mergers' partial effects on the retail prices of different parties, estimated under the initial assumption that the change in the firms' pricing policies all took place in January 2008.[9]

Table 1. Illustrative results for separate merger effects

	Agip/Esso	Lukoil/Jet
Own effect on buyer firm's stations	0	0.80%
Own effect on acquired firm's stations	0.70%	0
Competitor effect on stations in buyer's vicinity	0	0
Competitor effect on stations in acquirer's vicinity	0.50%	0

[6]Although there is no theoretical model backing this last result, this conjecture is based on the fact that it is usually the management of the buyer firm that takes over the business and pricing decisions of the acquired firm. We can also have more reason to believe that the buyer firm already had a more successful business strategy in place.

[7]Both mergers were of moderate size, the fifth and fourth biggest firm taking over the stations of the seventh and sixth firm, respectively.

[8]For this reason, our methodology cannot be used for the ex post evaluation of a merger affecting all (product or geographical) markets in the same way, or when there are too few distinct markets.

[9]The positive elements are parameters significant even at 1%, zero elements indicate parameters that are not significant at 5%.

Our first result is that neither merger contributed substantially to retail price increases, as all estimated price changes are less than one percent. Second, the two mergers had different effects on the merging firms depending on their role in the merger. These differences are broadly in line with the theoretical predictions. For the Agip/Esso merger, there are significant effects on the pricing of the acquired Esso stations and their competitors, and the price change is larger at Esso stations than at competitors' stations (although the difference is not significant). For the Lukoil/Jet merger, we also find that own effects are larger than competitor effects, but a significant effect is found only for the buying firm's stations. These different patterns of price effects for the two mergers may be explained by the difference in the merging firms' pricing policies and possible efficiency effects.

Unfortunately, the date when a merger effectively has an actual economic effect on the respective firms (the so-called effective merger date) is usually unknown to researchers, and changes in pricing policies might even be gradual. Therefore, great care should be taken when selecting the effective merger date to use in estimating the price effects of mergers, as this choice my have large impact on results. We apply several methods to show that the qualitative results discussed above emerge robustly when different effective merger dates are assumed. The magnitude of the estimated effects can change, but the estimated price changes due to the two mergers always remain negligible.

3.2. Overview of the relevant literature

The number of ex post merger evaluations (or so called merger retrospectives) has been growing considerably since 2000, partly because of an increased need to evaluate the performance of antitrust policies.[10] The principal statistical method used in these studies was difference-in-differences estimation, as is typical in program evaluation literature.[11] These ex post evaluations mostly studied transactions in industries with high merger activity, where price changes affected a wide range of consumers (and interested a lot of politicians and policy makers): hospital services,[12] airline ticketing,[13] banking products, basic consumer goods (typically food),[14] and gasoline.

Gasoline markets have always received attention, in particular during the recent years of large price changes, as fluctuations in petroleum prices were often followed by quick reactions in retail prices. Therefore, it was questioned whether the changes in wholesale conditions offered the only plausible explanation or whether certain anticompetitive practices also played a role.[15] A restructuring has taken place in many countries by a series of acquisitions on all supply levels, and therefore it is crucial to determine how the price effects of changes in wholesale conditions can be separated from changes in retail market structure.[16] Mergers involving companies with production facilities have always received more attention because of their ability to affect wholesale prices, but retail mergers are typically easier to analyze due to the availability of a larger amount of more transparent price data and variation in local market structure.

[10]Weinberg (2008) and Hunter et al (2008) provide two comprehensive reviews on ex post merger evaulations. Weinberg also discusses alternative methods to difference-in differences estimations.
[11]Imbens and Wooldrige (2009) give a detailed general review on the methodological problems arising in program evaluation.
[12]See Farrell et al (2009) for a recent overview on hospital merger retrospectives.
[13]See Armantier and Richard (2008).
[14]See Ashenfelter and Hosken (2008) for the ex post evaluation of five mergers in this sector.
[15]See for example the questions raised by the US Congress to the Federal Trade Commission in 2004. The summary of the FTC's view can be found at http://www.ftc.gov/opa/2004/07/gastest2.shtm
[16]The US Government Accountability Office reports 2600 mergers in the petroleum industry from 1990 till 2004. The GAO's econometric models analyzed the effects of the eight biggest transactions in detail. The report can be downloaded from http://www.gao.gov/products/GAO-04-96

A widely-cited paper by Hastings (2004) uses a simple difference-in-differences estimator to analyze how the acquisition of an independent station network by a branded network affected local retail prices in different geographic areas. She finds that removing an independent station raises retail prices significantly, but the increase in the share of branded (so called company-operated) stations alone does not explain higher prices. Hastings thus concludes that the identity of competitors is as important as their number in determining market conduct, which she interprets as support for a model with some product differentiation and brand loyalty on retail gasoline markets.[17] However, Hastings analyzes only the change in the pricing of competing stations to derive conclusions on the effect of the merger,[18] while theory suggests – and our paper also demonstrates – that the change in the pricing of acquired stations may be larger.

Taylor and Hosken (2007) use an approach similar to Hastings's in measuring the effect of a joint venture, but find no retail price increases resulting from the change in market structure.[19] The paper also illustrates some important implications for further ex post reviews: (1) it is more important to analyze the merger effects on retail prices than rack (wholesale) prices, (2) variation in gasoline supply should be taken into account whenever possible, and (3) the estimated effects can depend on the control regions used, therefore robustness checks are crucial when selecting the counterfactual.[20] The substantive difference between Taylor and Hosken's paper and

[17]This paper was criticized by Taylor et al (2010) both from a theoretical and empirical point of view. They failed to reproduce her results by using alternative data and also showed that her empirical result would not lead to unambiguous welfare effects in the underlying model she assumes.

[18]Not enough data was available on the acquired stations, as this was a random sample on prices in which minor brands and independent stations were underrepresented.

[19]Another merger concerning the same firm Marathon-Ashland was similarly analyzed by Simpson and Taylor (2008), and this study found no ex post evidence of a price increase either.

[20]Choné and Linnemer (2010) also use various local market definitions in order to find the robust ex post effect of a merger between two large parking companies.

ours is that while Taylor and Hosken examine the overall effect of the merger on city-level (average) prices, our paper takes a further step by separating the different effects a merger may have on various market players in each market.

The few studies analyzing the merger effects on rivals' prices used it to answer different research questions than we do. Kim and Singal (1993) find larger price effects for competitor airlines than for the merging airlines, which they attribute to merger-specific efficiencies passed on to consumers. In studying hospital mergers, Dafny (2009) argues that researchers should look at the effect on competitor prices particularly if it can be conjectured that the merger event might be correlated with the unobserved characteristics of the acquired hospitals, in order to avoid selection bias. However, as we have already mentioned, if there are no efficiencies realized by the merger, then these outsider effects provide only a lower bound for the insider effects.

3.3. Structural changes on the Hungarian retail gasoline market

The Hungarian retail gasoline market is moderately concentrated, with five main international oil companies (OMV, Shell, Agip (Eni), Lukoil and Hungary-based MOL) owning almost 75% of petrol stations and likely accounting for an even higher share of revenues. MOL's market share is the largest in terms of stations and it also has a leading role at the wholesale level with an upstream market share of at least 70%.

In 2007, Agip and Lukoil acquired all retail stations of the other two international oil companies present in Hungary, Esso (Exxon) and Jet (ConocoPhillips) respectively.[21] Both acquisitions were part of large transactions involving business

[21]The publicly available decisions can be downloaded from http://ec.europa.eu/competition/mergers/cases/decisions/m4723_20070724_20310_en.pdf

activities in multiple countries.[22] Table 2 summarizes the key facts concerning the two mergers.

Table 2. Summary of the two mergers analyzed

	Agip/Esso	Lukoil/Jet
Case number	COMP/M.4723	COMP/M.4532
Stations sold in countries	Belgium, Czech Republic, Finland, Hungary, Poland, Slovakia	Czech Republic, Hungary, Slovakia
Number of stations and market shares in Hungary	102 (9%) / 36 (3%)	42 (4%) / 30 (3%)
Transaction date	27/04/2007	18/12/2006
Notification date	19/06/2007	17/01/2007
Clearance date	24/07/2007	21/02/2007
First public sign of change in Hungary	Nov-07	Nov-07

Note that before receiving the clearance decision, the merging companies should act independently of each other. The change in business and pricing policies due to the merger usually takes place some months after the clearance, but this so-called effective merger date is not publicly known. The only publicly observable fact in our case is the date when the acquiring firms started repainting the acquired stations to their brand colors, which happened in November 2007 for both mergers.

Apart from the branded stations mentioned, the remaining 25% of stations are owned by a large number of small competitors. Only three chains had a larger than 1% market share based on the number of its stations in the relevant period of 2007-2008, and all three faced major changes during this time. The previously largest entrant Tesco continued to build new stations till the end of 2008.[23] The alliance of independent (white) stations Klub Petrol exited the market at the end of 2007

and http://ec.europa.eu/competition/mergers/cases/decisions/m4532_20070221_20310_en.pdf respectively.

[22]For this reason, the mergers were notified to the European Commission, which investigated and cleared them in quick Phase I investigations.

[23]Later in 2009, Tesco stations were acquired by Shell. This long-term lease agreement was cleared by the Hungarian Competition Authority (GVH), case number Vj-17/2009.

due to financial difficulties. Finally, a new alliance of independent (white) stations named Avia entered in 2008; several former Klub Petrol stations joined this alliance.

The evolution of the relevant firms' shares in station numbers is summarized in Table 3. The number of independent white pumps, indicated in the table as 'Other stations', is slightly uncertain, as their presence is not properly reported at the beginning of the observation period.

Table 3. Changes in station shares in 2007-2008

	Station shares 2007-01	Station shares 2008-12
MOL	29%	28%
Shell	16%	15%
OMV	14%	13%
Agip	9%	12%
Esso	3%	
Lukoil	4%	7%
Jet	3%	
Tesco	3%	4%
Klub Petrol	3%	0%
Avia	0	3%
Other stations	16%	18%
Total station number	1229	1335

3.4. Price data and stylized price developments

We analyze a panel database containing daily retail gasoline prices from the beginning of January 2007 till the end of December 2008. The database contains the price of 95-octane gasoline only, but 96% of gasoline sales are of this type. The source of our data is a public website helping consumers to compare gasoline prices: www.holtankoljak.hu (Where Should I Refuel?), run by a private company.[24] We analyze retail prices on Fridays, as the Hungarian wholesale price changes each

[24]The company conducting the price comparisons is independent of the retail firms, and is financed by online advertisements placed primarily by car manufacturers and insurance companies.

Wednesday morning,[25] and therefore most retail price changes occur on Wednesday and Thursday.[26] In total, we have 81253 price observations in 96 weeks for 1303 gas stations, more than 95% of stations in Hungary.[27] The panel is unbalanced, but there is no pattern in the missing data and the majority of the missings corresponds to the fringe white stations that are less relevant for our analysis.

In the observed period, the price of gasoline fluctuated between 230 and 310 Hungarian Forints (HUF), with an average of 281 HUF.[28] In order to filter out common shocks (particularly the change of the wholesale price), Graph 2 shows the differences between firm-specific average price and the national average price. Note that here the brand of the station refers to its original brand at the beginning of the observation period before either merger, so there is no composition effect in the changes.

These price differences show that the three largest firms (MOL, OMV and Shell) were able to maintain slightly higher prices than their competitors,[29] and Agip's prices gradually became closer to them in 2008. There is also a visible change in the pricing of Esso stations acquired by Agip in the middle of 2007, as their prices increased from the level of low-pricing firms to the national average. On the other

[25]Strictly speaking this is only the price change of the dominant wholesale company (Hungary-based firm MOL), but it supplies at least 70-80% of gasoline sold in Hungary. The change in the wholesale price is made public the previous Monday.

[26]Data gathered is based on the self-reporting of the stations by phone, and the biggest inquiry conducted is on Wednesday and Thursday. Therefore the Friday data are expected to be the most accurate, and we also have the most observations for Fridays.

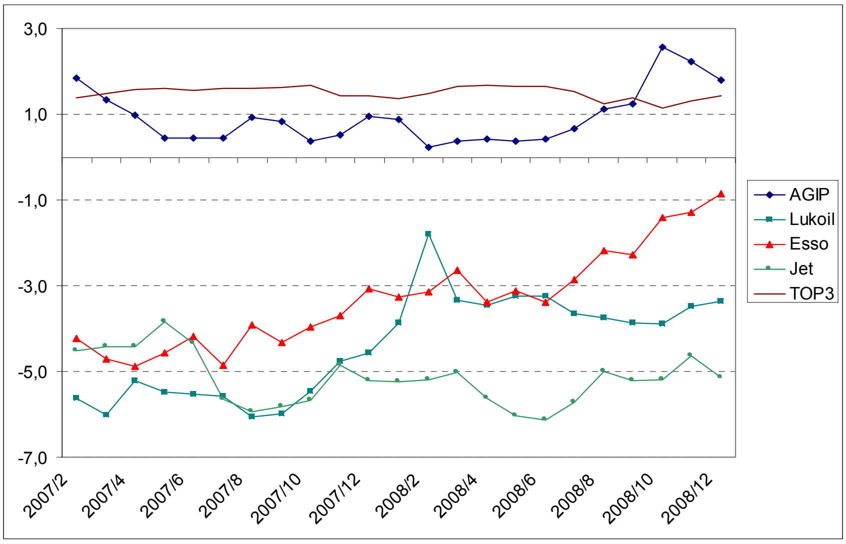
[27]The missing stations are all white stations or belong to small brands with few stations. Price data are not reported for 8 weeks, which was because of holiday periods and a shutdown problem of the website.

[28]The exchange rate also fluctuated during these two years, but one can make easy conversions with the approximation of 250 HUF = 1 Euro.

[29]The graph shows the average of the top 3 firms' prices together as no substantial differences can be observed between their average prices.

hand, Lukoil and Jet stations appear to have maintained their low-pricing policies, although Lukoil's prices increased slightly starting from the second part of 2007.

Graph 2. Monthly differences between firm-level and national average price



Although we cannot observe the individual costs of retail firms, we do have a good proxy for the wholesale price of gasoline we can use. Each Monday, the change in MOL's wholesale list price becomes publicly known, and as MOL serves most retailers and has significant market power upstream, we believe that the change of this wholesale price can serve as a good indicator of the change in marginal costs. Therefore, we will refer to MOL's wholesale list price as the wholesale price, and define the margin of a brand or station as the simple difference of the respective retail price and the wholesale price. The average margin varies between 10 and 18 HUF with an average of 15 HUF, which is about 5% of the retail price.

We should note that both the retail prices observed at the stations and the wholesale price are only list prices, while most retailers offer loyalty discounts in the form of loyalty cards or fleet programs. If larger retailers offer larger discounts,[30]

[30]This conjecture is hard to test correctly, because the discounts often do not take the form of a direct price decrease for gasoline, but for example bonus points (price discounts) for shop purchases.

the actual price differences between smaller and larger retailers might be smaller than shown in Graph 2. Similarly, at the wholesale level retailers receive individual discounts from the list price, which are not observed. However, assuming the size of these discounts remains stable over the observed period, these measurement errors are mostly taken care of by the difference-in-differences estimation method we use. Of course, it might be the case that a merged firm achieves a larger quantity discount due to increased sales or changes its consumer discount policy, but this change will be captured by our estimated merger-specific effects.

3.5. Local markets and characteristics of local competition

In order to analyze the effect of structural changes on local prices, we should first define the areas where a given station's pricing policy might constrain other stations' pricing, and therefore a structural change concerning the given station would have an effect on the other stations. This approach to delineating local markets is very similar to the usual first step in competition policy of "defining the relevant markets", where the competitive assessment should be carried out.[31]

In this paper, we use an economically reasonable proxy for local markets: the 168 statistical municipalities defined by the Hungarian Central Statistical Office (KSH). The statistical municipalities are delineated by various survey techniques as distinct geographical areas where inhabitants perform the majority of their social and economic activities (such as traveling, working and shopping), so it seems reasonable to assume that consumers shop around primarily in this area and retailers consider

[31]Note, however, that competition policy cases analyzing retail gasoline markets took a rather conservative approach by defining the relevant geographical market as national. It was only in a recent merger case in 2008 where DG Competition took the view that although the market is defined as national, the competitive assessment should take local aspects into account – see COMP/M.4723 StatoilHydro/ConocoPhillips decision §26-29, downloadable from http://ec.europa.eu/competition/mergers/cases/decisions/m4723_20070724_20310_en.pdf

the stations in the municipality as their main local competitors.[32] It might be of course the case that two competitors in the same local area do not exert the same degree of competitive pressure on a given station due to the varying distance between the stations, but further station-level controls in our estimations can partially take care of this problem. Similar proxies for local markets were used in several papers. Hortacsu and Syverson (2007) for example use the regional economic areas defined by the Bureau of Economic Analysis (BEA) to delineate cement markets in the USA; Focarelli and Panetta (2003) and Sapienza (2002) delineate Italian provinces as relevant geographic markets for bank deposits.

The shortcoming of the statistical municipality proxy is that it defines borders that separate markets. Competitors of a given station will be all other stations in the statistical municipality, but stations in another statistical municipality, no matter how close they are to this given station, will be not. This 'border effect' can lead to questionable classification of competitors near the borders of the statistical municipality and therefore some misclassification of treatment and control units.

An alternative way to delineate local markets would be to define a catchment area around each station, in which the given station provides a viable alternative to consumers visiting the other stations.[33] Note, however, that the choice of driving time / distance used in these delineations that defines how much overlap there will be among local markets remains arbitrary. Moreover, the stations falling in the same

[32]There are two further practical advantage of using statistical municipalities. First, the ZIP code of each station can be automatically linked to a municipality, which ensures that local markets do not overlap and their number can be kept at a tractable level. Second, the Statistical Office also discloses economic indicators (such as population, number of cars, taxable income) for each of them, which can be used to control for local differences in our estimations.

[33]For example, Hastings (2004) uses circles of a one-mile radius around each gasoline station, but estimates some of her results by using different radiuses. It is also possible to work with different measures of distances, like traveling time, which is more typical in analyzing supermarket mergers for example (see for example Ashenfelter et al (2006)).

catchment area are still assumed to exert the same competitive constraint, so the shortcomings of the previous approach are not completely solved.[34]

We do not include gasoline stations located on highways in our analysis, because substitution possibilities and therefore competitive conditions are markedly different at these stations.[35] The capital of Hungary (Budapest) with its 183 stations is defined as one statistical municipality, and therefore this outlier is also excluded.

Descriptive statistics of the remaining 167 statistical municipalities seem to indicate oligopolistic market structures with a few market players, which might signal that these municipalities constitute indeed a good approximation of local markets.[36] The average number of differently branded stations (major firms) in a local market is 3.2 (standard deviation 1.8), while the average number of stations in a local market is 6.3 (standard deviation 6.7).[37]

On top of the variance between the average characteristics of local markets, the major firms are also differently distributed among these markets, and it is this variance in market structure that we will heavily exploit. Table 4 and 5 illustrates the overlap of the merging firms in local markets and shows a different geographical pattern emerge for the two mergers to be analyzed.

[34] Note also that distance might not be the only source of horizontal differentiation between stations.

[35] In Hungary, highways can be entered only after paying the toll, and exits can be quite far from each other. Therefore it seems unlikely consumers would enter and exit the highway for the sake of a potentially lower pump price. The average price at petrol stations on highways is only 3-4% higher than at other stations, but we see a slightly different trend in highway prices than in off-highway prices. Because of these facts, it is usual to define gasoline stations on highways as different geographical markets.

[36] Bresnahan and Reiss (1991) studied oligopolistic markets with entry costs and analyzed the relationship between local market size and the number of sellers (see also Campbell and Hopenhayn (2005)), and they found similar distributions of firms in some industries to ours - see especially the distribution of automobile dealers in their Table 2, which probably has the closest connection to the gasoline market in their sample. We can also check in our case that the number of firms is strongly correlated with indicators proxying local market size (0.96 with taxable income, 0.97 with population).

[37] Note that firms (brands) can have multiple stations in some local markets (usually in the larger ones), and we include white stations as well when looking for the total number of stations.

Table 4. Number of local markets where Agip and Esso are present

	Esso present	Esso not present	Total
Agip present	12	52	64
Agip not present	4	96	87
Total	16	135	

Table 5. Number of local markets where Lukoil and Jet are present

	Jet present	Jet not present	Total
Lukoil present	6	32	38
Lukoil not present	11	115	113
Total	17	134	

Esso stations acquired by Agip were direct competitors of Agip in 75% of the local markets where Esso is present, while this overlap is only 35% for the Lukoil/Jet merger. On the other hand, the acquisition of Jet stations increased Lukoil's presence on local markets by almost 30%, but this expansion is only 6% for Agip. Therefore, the acquisition of Esso by Agip can be seen more as a merger with a direct competitor, while the Lukoil/Jet merger resulted more in market expansion. This may lead to a conjecture of a larger price effect resulting from the Agip/Esso merger, which may look consistent with the evolution of average prices at the firm-level presented on Graph 2.

However, the analysis of descriptive data can give only preliminary conjectures on merger effects. Definitive results can only be obtained by the thorough analysis of changes in local prices while controlling for other factors affecting the prices.

3.6. Estimation method and identification of ex post merger effects

3.6.1. Estimation of merger effects

Our main aim is to differentiate among eight types of price effects resulting from the two mergers, as demonstrated by Figure 1 in the Introduction. Naturally, we want to separate the price effects of the two mergers (Agip/Esso and Lukoil/Jet effects). However, theory suggest that price effects are heterogeneous and therefore we want to estimate four separate effects for each merger. First, we want to differentiate between the direct effects on the prices of the merging parties and the indirect effect on the prices of their respective competitors (own versus competitor effect). Second, we also want to separate the price effects associated with the two different parties in each merger (buyer and seller effects).

The eight merger effects are estimated with a difference-in-differences estimator. We chose the difference-in-differences estimator, because we have only few observables to account for differences in station characteristics and the double differencing removes the time invariant differences among stations affected and unaffected by the mergers. Moreover, it also controls for time trends that are not related to the mergers. Since we track stations over the sample period, the difference-in-differences estimator is implemented in a panel regression framework with station and time fixed effects. Our estimated equations take the following standard form:

$$\begin{aligned} y_{it} &= \sum_{j \in Agip, Esso, Lukoil, Jest} \alpha_{1j} own_{jit} + \\ &+ \sum_{j \in Agip, Esso, Lukoil, Jest} \alpha_{2j} competitor_{jit} + \\ &+ \sum_{k} \beta_k control_{kit} + u_i + v_t + \varepsilon_{it}, \end{aligned} \tag{3.1}$$

where α-s and β-s are parameters to be estimated, i indexes stations, t indexes time and j indexes the merging parties. The dependent variable y_{it} is the outcome variable of interest, which can be the price or the price-cost difference in absolute (margin) or in relative terms (markup). Dummy variables own_{jit} capture the merger effect on merging parties and take the value of one after the merger for the stations of each merging firm and zero otherwise. Variables $competitor_{jit}$ capture the merger effect on competitors and take the value of one if station i has merging party j as a competitor after the merger and zero otherwise. Depending on additional data possibilities and considerations, we can add further variables (summarized now in $control_{kit}$) to the model to control for other, possibly time variant factors that can affect price changes. Finally, the error component u_i is the station fixed effect, v_t is the time fixed effect and ε_{it} is the disturbance term.

The inclusion of station and time fixed effects basically transform the data into differences from the respective means. This ensures that the parameters of the own_{jit} and $competitor_{jit}$ indicator variables are indeed the difference-in-difference estimates: they capture the difference in the margins between stations affected by the merger and take their difference before and after the merger. Moreover, station fixed effects control for unobserved heterogeneity, which is important in our case because we do not fully observe all relevant characteristics of different stations and the size of local demand for gasoline. Time fixed effects control for changes in common unobservable variables to all stations in a given period.

3.6.2. Identification of the eight merger effects: treatment and control groups

Although the specification in (3.1) is quite standard, it is worthwhile to discuss identification of the merger effects a bit more in detail. The main question is whether the estimated parameters can be interpreted as the causal effects of the merger on firms' margins. We argue that this is the case if the merger effects can be interpreted as average treatment effects with the merger being the treatment. In order to see whether such interpretation is valid, we have to clarify how the merger effects in (3.1) compare the margins of merging stations and their competitors to margins of all other stations.

A merger can be interpreted as a natural experiment, because the stations' individual prices are usually set at firms' headquarters, so it is reasonable to assume that the merger changes the pricing policy of the firm itself. This policy change affects stations of merging parties and their competitors after the merger date, while it does not influence all other stations' behavior. We can then estimate the average change in the realized station-level prices after the adjustment in firms' behavior took place.[38]

Any natural experiment defines a treatment group, which in the present case is the set of stations affected by the merger. Obviously stations belonging to the four merging parties were affected by the merger and therefore belong to the treatment group. However, through the price equilibrium, also competitors are affected by the merger and should form part of the treatment group. Defining competitors of merging stations is not straightforward. Our choice is to define them as the stations that are present in the same local market as any of the merging parties. This definition implies that the treatment group is effectively defined over the local markets. Following from the treatment group definition the control group is the set of stations that are located in markets where no merging station is present.

Natural experiments can be exogenous or endogenuous. In the merger context the first implies that the price outcomes of stations before and after the merger

[38]A structural model of horizontal differentiation could for example lead to an equilibrium pricing condition for a firm that sets a uniform price P with the condition that any station should decrease its local price by X if firm A is present and Y if firm B is present. The merger can change the parameters in this equilibrium price setting rule to X' and Y', which could imply different price changes for two stations of the same firm, as they face different competitors on their local markets.

are independent of the merger decisions themselves, i.e. there is no selection bias. The Hungarian acquisitions formed only small parts of larger transactions involving stations in multiple countries, and all Esso and Jet stations in Hungary were sold to their respective buyers. Both factors considerably decrease the chance for a selection bias. This is an advantage compared to other studies that observe mergers initiated in the observed market. The difference-in-differences estimator can be unbiased even if the merger decision is endogenuous (not independent of the post merger outcomes given the observables), but exogenous merger decisions make our estimation potentially more robust.

The difference-in-differences estimator compares the average margin of the treatment group (stations taking part in the merger and their competitors) and the control group (all other stations) before and after the merger. The motivation behind the difference-in-differences comparison is that observations on unaffected stations (the control group) can form a counterfactual by informing us about what would have happened to the merged stations had the merger not taken place. Controlling for additional factors like demand and market structures ensures that we compare as similar subjects as possible.

Since the treatment group is based on the different presence of merging stations in local markets, the treatment group will be heterogenous. For example, there will be local markets where only Esso is present, markets where both Esso and Agip are present and also markets where Esso and Lukoil are present etc. Overall there are 16 possible presence combinations, which implies 16 different treatment group types. The central question of identification is how price information on stations in these different treatment types are combined to produce the eight merger effect estimates specified in (3.1).

Table 6. Control / treatment groups identifying Esso own and competitor effects

Presence combination (AELJ)	Number of local markets	Number of station-week observations					
		Agip	Esso	Lukoil	Jet	Controls / Non-merging competitors	Total
0 0 0 0	75	0	0	0	0	20,191	20,191
1 0 0 0	37	3,866	0	0	0	13,955	17,821
0 1 0 0	3	0	130	0	0	1,558	1,688
0 0 1 0	16	0	0	1,571	0	7,694	9,265
0 0 0 1	4	0	0	0	290	2,625	2,915
1 1 0 0	4	338	224	0	0	2,374	2,936
1 0 1 0	10	1,006	0	1,040	0	4,244	6,290
1 0 0 1	3	398	0	0	364	3,149	3,911
0 1 1 0	0	0	0	0	0	0	0
0 1 0 1	1	0	124	0	40	587	751
0 0 1 1	1	0	0	86	31	489	606
1 1 1 0	2	241	221	218	0	2,075	2,755
1 1 0 1	3	373	211	0	266	3,670	4,520
1 0 1 1	2	432	0	331	215	2,367	3,345
0 1 1 1	0	0	0	0	0	0	0
1 1 1 1	3	652	204	349	275	4,227	5,707

Out of the 16 possible treatment types we observe only 14 in the sample as demonstrated by Table 6. In this table the 4-digit code indicates the respective presence of Agip, Esso, Lukoil and Jet stations in a local market, where 1 indicates the presence of the respective firm and 0 its absence. For example, in treatment type 1010 Agip and Lukoil are present with Esso and Jet absent, while in treatment type 1110 one or more Esso stations are present as well. For each treatment type Table 6 shows the number of local markets with such treatments, the number of station-week observations for each merging brand and the number of station-week observations for non-merging competitors in those markets. For example the treatment type 1010 is found in 10 local markets and contains.1006 observations on Agip stations, 1040 observations about Lukoil stations and 4244 observations about other non-merging competitors. The control group (0000) is indicated in the first line of the table with 75 markets and 20191 observations. Most treatment types are less frequent, however, and are observed in less than five local markets. Only three treatment types with

Agip and Lukoil presence consist of ten or more markets. The median treatment type contains around 3000 station-week observations.

Most of the treatment observations belong to competitors (median treatment type contains 2374 cometitors).[39] In order to see clearly how many observations there are in total on the different merging parties, Table 7 summarizes the number of markets and station-week observations of each merger effect. As one can see, the number of observations on smaller merging parties is still adequate to estimate the merger effects. In Table 7 the competitor stations are double counted: in a market with both Agip and Lukoil presence all other station are added both to Agip and Lukoil competitor categories. This allows one to see exactly how many competitor observations each merging party has in total.

Table 7. Treatment and control groups identifying all own and competitor effects

	Treatment groups		
	No of local markets	No of stations-week observations	
		Own stations	Competitor stations
Agip	59	7306	39979
Esso	16	1114	17243
Lukoil	34	3595	24373
Jet	17	1481	20274

The difference-in differences estimator as specified in (3.1) estimates the eight merger effects based on the double difference of the treatment group observations and the control group observations before and after the merger dates. For example, the estimate of the Esso own effect combines the margin difference of Esso stations in treatment group 0100 and the control group with the difference of Esso stations

[39]The much larger number of observations to identify competitor effects might lead researchers to rely on them more in estimating merger effects (see for instance Hastings (2004)), but as we will see from our results, competitor effects are only a lower bound for own effects.

in treatment group 0110 and the control group. In fact, for each merger effect there are eight possible presence combinations (treatment group types) that are used for identification.

As long as one estimates only the own effects[40] such pooling of different treatment types can be allowed to estimate a reduced number of merger effects. The merger effect estimates can still be interpreted as the average treatment effect of the mergers. However, including competitor effects implies that in treatment group 0110 the Esso station receives a double treatment: one from its own merger and one from being the competitor of a Lukoil station that also participated in a merger. This violates the Stable-Unit-Treatment-Value-Assumption of the quasi-experimental design since treatments received by another unit (station) influence the outcome of another unit (station). As a result, these two effects cannot be identified separately in general. Nonetheless, competitor effects are important, because they express the simple fact that stations interact with each other through the market. The specification in (3.1) circumvents this identification problem by imposing an additional assumption: both own and competitor merger effects are linearly additive. This additional assumption allows one to use also the multiple presence treatment types to identify the merger effects, given that the single presence treatment types identify the merger effects even without this assumption. Additiv linearity is a quite strong assumption, however. Therefore, among the robustness checks we also include an alternative merger effect definition where such treatments are excluded from the identification.

[40]This is the case for example in Prager and Hannan (1998) who compare interest rates of banks in regions where both merging firms compete to regions where they do not, or in Vita and Saches (2001) who compare hospitals with a similar number of beds, size and location differing only in whether they were present in a county where the merger occurred or not.

A matching estimator could be an alternative to the difference-in-differences method. Such an estimator looks attractive if one looks at the normalized differences of covariates between the treatment groups and the control group. As Table 8 demonstrates, these differences generally exceed one quarter, which suggests poor covariate overlap among control and treatment groups.[41] Unfortunately, we do not have sufficiently rich station characteristics for proper matching, and the difference-in-differences estimator can overcome this shortcoming of our dataset by the double differencing. Of course, poor overlap is not good news even from the difference-in-differences perspective: the parallel trends condition may not be met. Panel data, however, allows to test for some restrictions imposed by the difference-in-differences estimator: one can compare the pre-treatment averages of control and treatment groups, which should not differ statistically.

Another argument against a matching estimator is that the unconfoundedness or the selection on observables condition, which is required for this estimator, is most likely not met. Although the mergers are exogenous to the Hungarian market, the treatment groups (local markets) are not selected at random: stations are located in markets based on pervious optimizing decisions of firms. There are many aspects of these location decisions that are unobserved. Margins after the merger can be functions of these unobservables that also affect the assignment to treatment status. For example, based on information unobservable to us, Esso could have bought a station at a good location knowing that it can earn high markups by doing so. Therefore, a high markup station is selected to be an Esso station and part of the treatment based on unobservables. Consequently, assignment to treatment

[41]Imbens and Wooldridge (2009) suggest this rule of thumb to evaluate the normalized differences.

status and outcomes in terms of the margin are not independent. This violates the matching estimator's assumptions.

Table 8: Normalized difference of covariates from control group

Treatment	Number of competitors	Cars per capita	Taxbase per capita
Agip own	0.86	0.67	0.66
Esso own	1.28	1.15	1.67
Lukoil own	0.52	0.46	0.23
Jet own	0.99	1.16	1.41
Agip competitor	1.04	0.98	1.14
Esso competitor	1.25	1.10	1.44
Lukoil competitor	1.22	1.00	1.19
Jet competitor	1.19	1.10	1.36

A final issue to clarify is the sensitivity of results to the specific market definition, since treatment and control unit definitions depend on the definition of local markets. A catchment area type market definition would clearly categorize stations into treatment and control units somewhat differently. Catchment areas are usually defined to be smaller than our statistical municipality proxy, which would imply the following differences in control and treatment unit definitions. First of all, using a catchment area definition, even within the same statistical municipality there would be control units - stations that do not belong to any of the distance based local markets. This would, ceteris paribus, increase the number of control units with observations that are potentially affected by the merger, which would reduce the difference between control and treatment units. Second, along the borders of two statistical municipalities - one treated and one non-treated municipality - control stations from the non-treated municipality would be redefined as treated. This would reduce the number of control stations with units that are most likely affected by the merger. As a consequence one would expect the treatment-control difference to increase. These two definitional changes move estimates into opposite directions and the end result is ambiguous. Third, there would be much more competitors belonging to only one merging station. This reclassification of competitors could potentially have the largest effect on the estimates, although its overall effect is unclear. These three differences in control and treatment unit definitions show that it is not clear-cut which market definition is better.

3.7. Estimated ex post merger effects by assuming a known treatment date

3.7.1. Baseline estimation

When we estimate the price effects of mergers in this Section, we make an assumption on the exact time when the two mergers started to affect the pricing of each firm, and that it takes place as one discrete change with immediate effect for all firms. We set this so-called effective merger date at the first day of January 2008 for both mergers, implying that we have exactly one year before the treatment and after the treatment. In the next Section, we discuss how to relax this assumption and the effect it has on our results, and also explain why January 2008 can be a good candidate for both effective merger dates.

We estimate the following version of Equation (1):

$$\begin{aligned} y_{it} = & \sum_{j \in Agip, Esso, Lukoil, Jest} \alpha_{1j} own_{jit} + \\ & + \sum_{j \in Agip, Esso, Lukoil, Jest} \alpha_{2j} competitor_{jit} + \\ & \beta_0 + \beta_1 size_{it} + \beta_2 size_{it}^2 + \sum_{k=3}^{12} \beta_{3j} type_{it} + cm_{it} + u_i + v_t + \varepsilon_{it} \end{aligned} \tag{3.2}$$

Although our main variable of interest is the price and the price effects, we will use the margin - the absolute difference of retail price and cost (wholesale price) - as our dependent variable. The main reason to do so is that prices are non-stationary while margins are, but we will show that running estimations on prices and controlling for costs does not change our results considerably. However, with the margin as a the dependent variable, the estimated values of α can be still interpreted as price effects if the wholesale price does not change due to the merger. Even if the merger does affect the firm-specific wholesale price (due to a volume discount for example),[42] the change in costs is passed on almost completely to retail prices (as our later estimations will show), so it can be interpreted as a price effect.

We take into account changes in market structure other than the two mergers by including the number of stations ($size_{it}$ and also its square) and indicator variables for the ten largest firms' presence in the same local market ($type_{it}$) in the regressions. The main condition for the validity of the difference-in-differences estimator is that there should be no differential trend among control and treatment stations. In order to control more strictly for common time variation we also use county-time fixed effects (cm_{it}).[43]

Table 9 shows our estimation results. Of the control variables, we only provide results for those that we interpret later. [44]

[42]As we do not observe the firm-specific wholesale price at which the retailer buys the gasoline it resells, but only a publicly observable proxy, we cannot test this hypothesis.

[43]There are 19 counties in Hungary, each containing 8 local markets (municipalities) on average.

[44]The estimated controls for the presence of merging firms are not significant, so there is no need to correct the estimated merger effects accordingly.

Table 9. Results of estimating Equation (2)

Dependent variable	Margin (in HUF)
Merger effects	
Agip own	0.10
Agip competitor	0.15
Esso own	1.78**
Esso competitor	1.24**
Lukoil own	1.95**
Lukoil competitor	0.20
Jet own	-0.43
Jet competitor	0.00
Controls (selected)	
Tesco	-1.04**
Klub Petrol	-0.71**
Avia	0.48**
No of stations	0.08
No of stations squared	-0.002**
Number of observations	82701
Within R2	0.27

* significant at 5% level, ** significant at 1% level

The estimated merger effects are only significant for Lukoil and Esso. In order to see the magnitude of these price effects, we show the relative change in retail prices and firm-level margins (in parentheses) in Table 10.

Table 10. Relative price (margin) changes due to different merger effects

	Agip/Esso	Lukoil/Jet
Own effect on buyer firm's stations	0	+0.8% (+13.1%)
Own effect on acquired firm's stations	+0.7% (+11.4%)	0
Competitor effect on stations in buyer's vicinity	0	0
Competitor effect on stations in acquirer's vicinity	+0.5% (+8.7%)	0

The results show that both mergers had a positive but non-substantial effect on retail prices, as all significant effects are less than 2 HUF that is less than 1% of the average price. In terms of margins, however, the mergers provided a substantial change in Esso and Lukoil stations' margins and also for stations in Esso's vicinity.

In line with our initial expectations, we see that both mergers resulted in different effects on the various firms, depending on their role in the merger. The Agip/Esso merger increased the prices of the seller's stations and also of their competitors, but the Lukoil/Jet merger had a positive effect on the prices of the buyer's stations only. Concerning the main theoretical prediction, the own effect is indeed significantly larger from the competitor effect in the case of the Lukoil/Jet merger. For the Agip/Esso merger, the point estimates also indicate a larger own effect than competitor effect, but this difference is not statistically significant.

We can provide a possible economic interpretation of the two transactions that may also explain why different patterns emerge in the price effects of the two mergers. In the case of the Agip/Esso merger, it is likely that the acquisition of Esso did not change the potential competitive pressure on Agip from low-pricing brands,[45] which can support why the Agip effects are not significant. On the other hand, the Esso stations became part of a larger firm with a reputation for higher quality, so the price increase on Esso stations (and therefore of its local competitors) can be likely attributed to an upwards brand repositioning of the Esso stations.

The acquisition of Jet expanded Lukoil's presence and recognition considerably, providing a plausible explanation for the increase in Lukoil prices.[46] However, no significant effect is found on the Jet stations, despite both firms having a similar pricing profile before the merger. A possible explanation could be that as Lukoil is a vertically integrated company, the marginal cost of the Jet stations could have

[45]As it can be computed from Table 6 listing the treatment and control groups, the low-pricing brands Lukoil or Jet are also present in 8 of the 12 markets where both Agip and Esso are present.
[46]Lukoil competitor effects are significant at the 10% level, but are non-substantial.

decreased because of the emerging self-supply opportunities. Therefore, an efficiency effect might have cancelled the otherwise positive price effect on Jet.[47]

The difference-in-differences approach also allows us to interpret some of the control variables as the effects of actual entry and exit of stations during the observed period.[48] The parameter estimate of the Tesco dummy for example indicates that the entry of a Tesco station to a given local market decreased prices by 1 HUF. The parameter estimate for Klub Petrol should be interpreted as the exit of a Klub Petrol station increasing prices by 0.7 HUF. Surprisingly, the entry of an Avia station increased prices by 0.5 HUF, but as a good part of bankrupt Klub Petrol's stations joined the Avia alliance, they likely could not sustain the very low prices they charged before.[49] A change in the number of stations in a given local market does not have a substantial price effect either (the squared variable is statistically significant, but very small).[50]

3.7.2. Separating the effect of local competitive interactions between merging parties

An additional important theoretical prediction to consider is that a merger may lead to larger price effects on those local markets where the two merging firms' stations were direct competitors of each other before the merger compared to those

[47] Unfortunately, our data does not allow to separate efficiency effects as we do not observe firm-level costs (input price).

[48] Such methods are usually called event or shock analysis. Ashenfelter et al (2007) discuss the pros and cons of using these techniques by presenting the econometric methods used in the famous Office/Staples merger. For a more general overview, see Davis and Garces (2010, Chapter 5).

[49] When they were still active, the average price of Klub Petrol stations was 7 HUF below the average national level (Lukoil's and Jet's average prices were about 5 HUF below national average).

[50] In this price-concentration relationship, theory predicts that the entry of a new firm has a negative effect on price, but this price effect is smaller in absolute value when there are more firms on the market. So the parameter of station number is expected to be negative, but the parameter of the squared number is expected to be positive.

where only one was present. Testing this hypothesis requires the separation of yet another set of effects: eight effects if the respective firm is without its merging party ("alone") in the local market, and eight effects if both merging firms are present in the local market.

Table 6 shows that there is enough variation in the composition of merging firms in distinct local markets for the formation of different treatment-control group pairs to identify each effect. We then estimate Equation (2) with sixteen treatments and report the results in the first two columns of Table 11. The third column of Table 10 contains the effects without taking into account the sole or joint presence of merging firms (our first estimates from Table 8), which are naturally the weighted averages of parameters in the first and the second column.

Table 11. Estimation results separated by sole/joint presence of merging firms

	Effect if present alone	Effect if both present	Overall effect
Agip own	0.23	-0.41	0,10
Agip competitor	0.26	0.00	0,15
Esso own	2.91**	1,25**	1,78**
Esso competitor	1.57**	1.15**	1,24**
Lukoil own	1.99**	1.54**	1,95**
Lukoil competitor	0.26	0.00	0,20
Jet own	-1.03	0.71	-0,43
Jet competitor	0.17	0.01	0,00

* significant at 5% level, ** significant at 1% level

Our first result is that the set of significant effects does not change if we separate the effects by the sole and joint presence of the merging firms (Esso own and competitor effects and Lukoil own effect). Second, the parameter estimates for the respective effects are in most cases not significantly different from each other statistically depending on whether only one or both merging firms are present. The exception being the Esso own effects and Jet own effects (the latter only at 10%).

In the case of the Esso own effect, a significantly larger price effect can be observed for those Esso stations with no Agip stations in the same local market. While this result may seem counterintuitive at first, it can support our previous discussion that the price increase at Esso stations was not caused by the elimination of a previously existing competitive pressure between Agip and Esso. On the local markets where Agip was not present before, the larger change may well have been due to the emergence of a more recognized brand.

In the case of Jet own effects, the point estimates show a negative effect on those markets where a Lukoil station was also present and a positive effect where there was not. This weakly significant difference is also consistent with the conjecture presented before that while the merger removed some competitive pressure exerted by Lukoil, Jet stations without the presence of Lukoil could have decreased their prices. Overall, the positive price effect was offset by an opposite effect that can be attributed to efficiencies.

To sum up, the separation of effects based on the sole and joint presence of merging firms in local markets does not bring strong evidence for the "significant lessening of competition". This is consistent with our previous result of non-substantial price effects and the fact that the Hungarian mergers only formed parts of large international transactions, and were not necessarily aimed to take over a strong local competitor.

3.7.3. Estimating alternative specifications

We now check whether alternative specifications of our estimated equation substantially modify our results. Table 12 shows the estimates of the eight merger effects for six different specifications.

The first three columns demonstrate how additional control variables affect our results. Specification (I) is the basic form of Equation (1) with only the necessary treatment dummies and standard cross-section and time fixed effects. In specification (II), we add controls for local competitors (number and type of rival stations), and also add time-county fixed effects in specification (III), which is the previously estimated and discussed Equation (2). These results demonstrate that the qualitative results do not change substantially, the statistically significant parameters are the Esso own and competitor effects and the Lukoil own effect. As the added controls are significant and the estimates provide economically sensible results, we include both sets of controls in all of our subsequent estimations.

Table 12. Estimated effects for various specifications of Equation (1)

	(I)	(II)	(III)	(IV)	(V)	(VI)
Dependent variable	Margin	Margin	Margin	Price	Log(Price)	Markup
Agip own	0,03	0,20	0,10	010	0,0002	0,0003
Agip competitor	-0,01	0,15	0,15	0,15	0,0004	0,0004
Esso own	1,49**	1,10*	1,78**	1,78**	0,0080**	0,0082**
Esso competitor	1,05**	0,67**	1,24**	1,24**	0,0052**	0,0054**
Lukoil own	1,98**	2,04**	1,95**	1,95**	0,0082**	0,0086**
Lukoil competitor	0,19	0,20	0,20	0,20	0,0009	0,0010
Jet own	-0,04	-0,16	-0,43	-0,43	-0,0003	-0,0004
Jet competitor	0,39*	0,30	0,00	0,00	0,0004	0,0004
Competitor controls	No	Yes	Yes	Yes	Yes	Yes
Time-county fixed effects	No	No	Yes	Yes	Yes	Yes
Within R2	0,17	0,19	0,27	0,99	0,99	0,46

* significant at 5% level, ** significant at 1% level

The last three columns demonstrate the results of estimating the merger effects on other dependent variables: the price and the markup. As neither of these variables are stationary, we should treat these estimation results with reservations. It is worth noting, however, that the regressions on price (specification IV) and the logarithm of price (specification V) produce similar results to estimates from our Equation (2). In these specifications, we include the cost (wholesale price) on the right-hand side as a control.[51] Specification (VI) shows the change in the markup (the price-cost difference divided by the price, also called the Lerner-index). This estimation is also in line with the margin estimate from Equation (2): a 0.8 percentage points change in the markup for Esso station corresponds to an increase of 13%.[52]

3.7.4. Alternative merger effect definitions

While discussing identification we pointed out that our specification assumes that merger effects are linearly additive so that all treatment observations can be used for the estimation. However, this assumption is quite strong. Therefore, we also present the results with alternative merger effect definitions. Table 13 presents the estimates where own effects are estimated only from markets where stations from one or the other merger were present (markets 1000, 1100, 0100 for the Agip-Esso merger and 0010, 0011, 0001 for the Lukoil-Jet merger) and competitor effects are estimated only from markets where a single merging firm was present. This way merger effects satisfy the conditions of quasi-experimental comparison without any further assumptions. We can use double presence markets for the identification of own effects, because after the merger stations of the merging parties do not compete with each other, therefore there is no competitor effect to be taken into account.

[51] In specification (IV), the estimated cost parameter is not significantly different from 1, which signals an (almost) complete pass-through at the retail level.

[52] The average markup is about 6% at the retail level.

Table 13. Alternative merger effect definitions

Dependent variable	Margin (HUF)
Merger effects	
Agip own	0.41*
Agip competitor	0.32*
Esso own	1.45*
Esso competitor	1.15**
Lukoil own	1.91***
Lukoil competitor	0.31
Jet own	-1.34
Jet competitor	0.60
Number of observations	82701
Within R2	0.27

* significant at 10% level, ** significant at 5% level, *** significant at 1% level

The results do not change significantly using the alternative merger effect definitions. The magnitudes are somewhat smaller for the Esso and Lukoil point estimates and they are also less precise (except for the Lukoil own effect). The only important difference is that the Agip effects are also statistically significant at the 10% level in this specification. The magnitudes of these merger effects are, nonetheless, very small. All in all these findings suggest that the additive linearity assumption holds in our case.

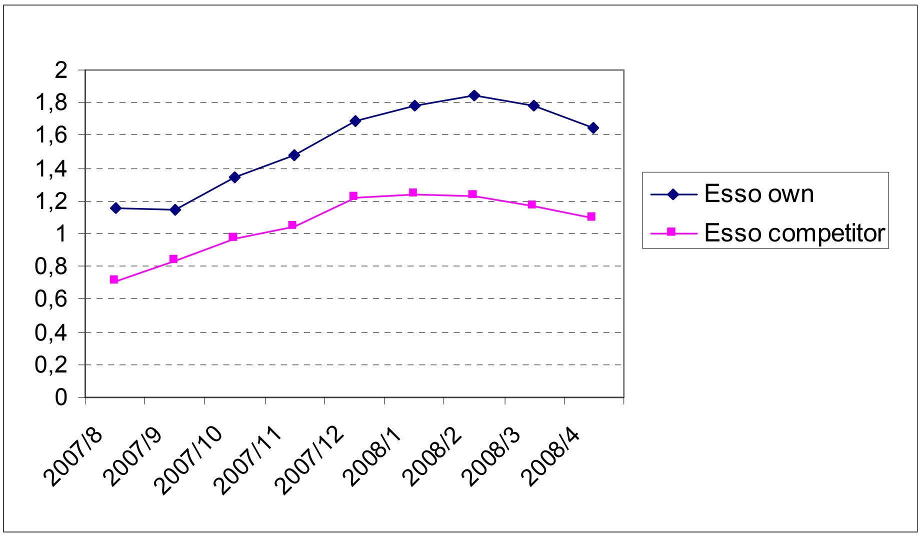
3.8. Sensitivity to the effective merger dates

Up to this point, we made strong assumptions on the effect mechanism of the mergers. First, we assumed that the effective merger date – that is the date the merging firms actually change their business and pricing policies – can be observed by the researcher, which is hardly ever the case. Typically, the only public information concerns the clearance date of the merger after which the merging firms are allowed to coordinate their business policies, but the change in firm's actual pricing may take several months. Some observable information may be available on firing managers or on rebranding decisions, but these are imperfect proxies of the effective merger date.[53] Second, we assumed that the change in pricing policy is a sudden discrete jump that is simultaneously happening for all firms, but adjustments could be gradual and competitors may not instantaneously react. It is therefore crucial to test the sensitivity of our results with regards to these assumption.

Initially, we analyze only the case when the two mergers takes effect at unobservable and potentially different dates, yet the effects are immediate. In our application, we estimate the merger effects in Equation (2) by gradually changing the effective merger date month-by-month from the first week of August 2007 till April 2008 for the Agip/Esso merger, and from the first week of March 2007 till March 2008 for the Lukoil/Jet merger.[54]

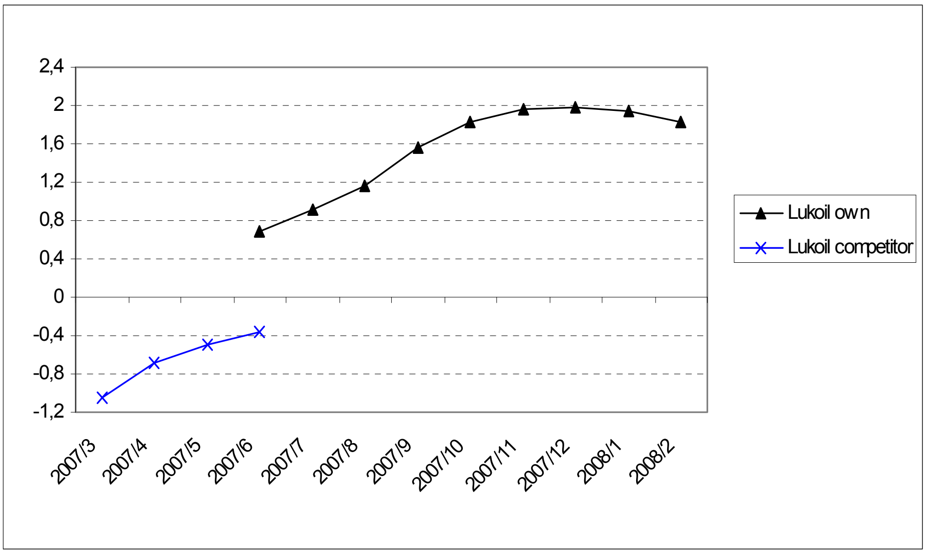
Graphs 3 and 4 show the own and competitor effects for Agip/Esso and then Lukoil/Jet for a range of effective merger dates, keeping the respective other effective merger date constant at January 2008.[55] We show only those estimates that are significant at 5%.

Graph 3. Significant Agip and Esso results for Agip/Esso effective merger dates



[53]In our case, the only information is that both Agip and Lukoil started to repaint the acquired stations in November 2007, but the brand of some stations was changed only several months later.
[54]Agip/Esso was cleared at 24-07-2007, Lukoil/Jet at 17-02-2007.
[55]Changing the Agip/Esso effective merger dates does not change substantially the estimated effects for Lukoil and Jet and vice versa, so we do not show results for all possible pairs of effective merger dates.

Graph 4. Significant Lukoil and Jet results for Lukoil/Jet effective merger dates



In the case of the Agip/Esso merger, the own and competitor effects are always significant and positive for Esso, but never significant for Agip.[56] For this merger, the choice of the effective merger date does not change the qualitative results, but the estimated effects almost double if the merger is assumed to make an effect six months after clearance rather than immediately. In the case of the Lukoil/Jet merger, however, we find only negative Lukoil competitor effects with effective merger dates before June 2007 and only positive Lukoil own effects with effective merger dates after this time. Therefore, for the second merger we may reach different qualitative conclusions if the effective merger date lies close to the clearance date, and the quantitative changes can be higher as well (the Lukoil own effect almost triples if effective merger date is December 2007 instead of June 2007). We see that all effects increase if we start to move the effective merger date from the clearance date, and the point estimates reach their peaks between December 2007 and February 2008.

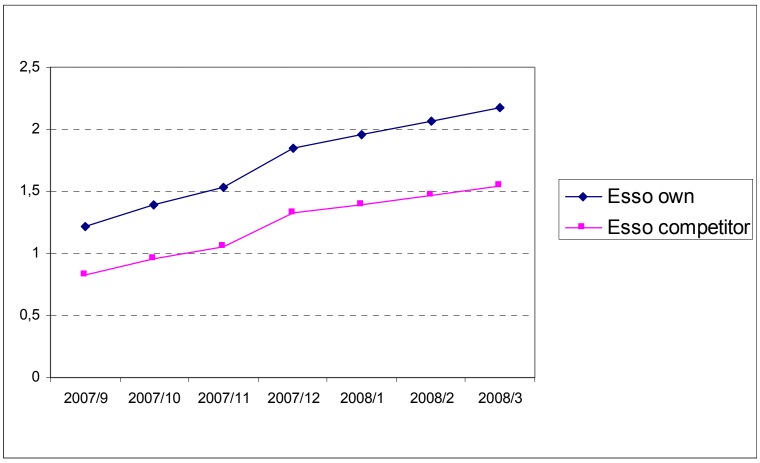
We may use the above results for a speculative reasoning on the effective merger date. If the treatment indeed causes a discrete and immediate price increase and there are no

[56]The own effect is also always larger than the competitor effect, although the difference is not statistically significant.

other shocks affecting the industry, then the estimates for the assumed effective merger dates should be increasing before the actual effective merger date and decreasing after. Therefore, if we observe a (statistically significant) peak in the pattern of estimated merger effects then the date of this peak can be a candidate for the effective merger date. This is what we have done in Section 7 by pinning down both effective merger dates to January 2008 and even these estimates showed negligible price effects for both mergers. Note that one should be cautious with this approach of selecting the effective date, but we can use the results to find an upper bound for the effects of a merger.

We now turn to the second potential issue, namely that the change due to the merger can be gradual, because of lengthy adjustment periods or differences in firms' reaction time. In this case, some observations fall in an intermediate period, and it might be beneficial to exclude this "window" period from the before-after comparisons.[57] In our application, we estimate Equation (2) by excluding a window starting from September 2007 for the Agip/Esso merger and July 2007 for the Lukoil/Jet merger,[58] and change the end of the window month-by-month till March 2008.

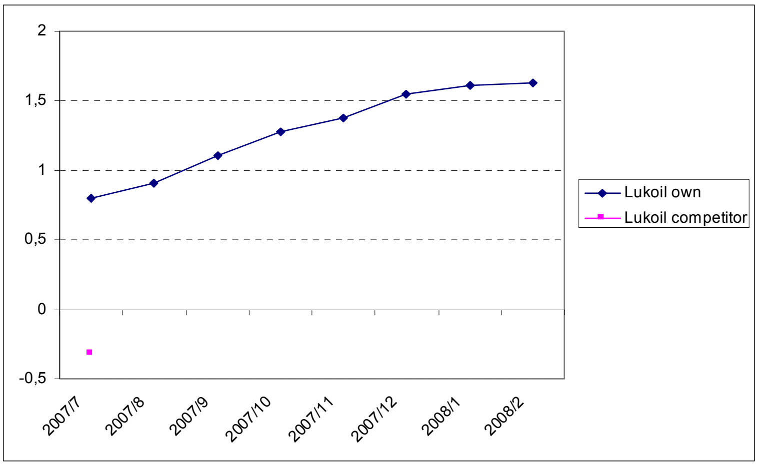
Graph 5 Significant effects for different Agip/Esso window end dates



[57]For the same reasons, Ashenfelter and Hosken (2008) estimate merger effects by leaving out observations within 3 months of the clearance date.

[58]If we start the Lukoil/Jet window before June 2007, only the Lukoil competitor effects will be significant.

Graph 6. Significant effects for different Lukoil/Jet window end dates



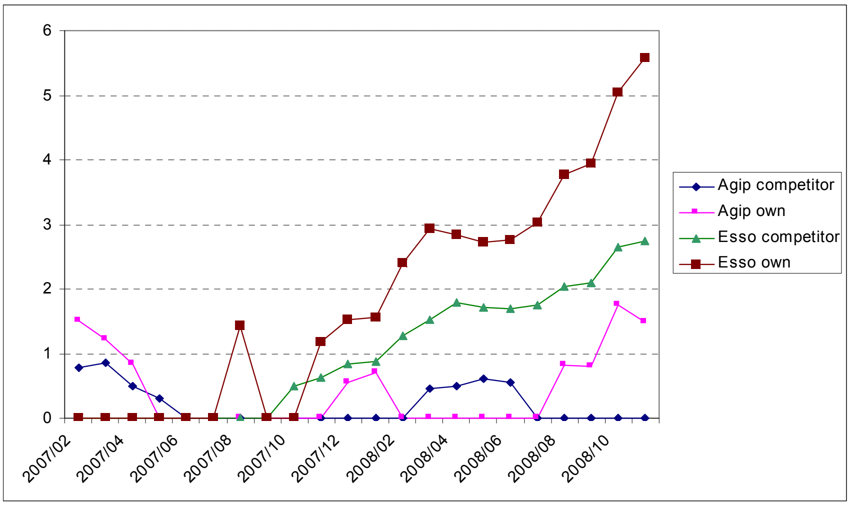
Graphs 5 and 6 show the significant effects (at 5%) of the respective merger for different window end dates, by keeping the other effective merger date at January 2008. The qualitative results do not change significantly on whether we include a window period or not: the set of significant effects remain the same and the effects are always larger when fewer observations from 2007 belong to the "after treatment" regime.

3.9. Control-treatment margin differences

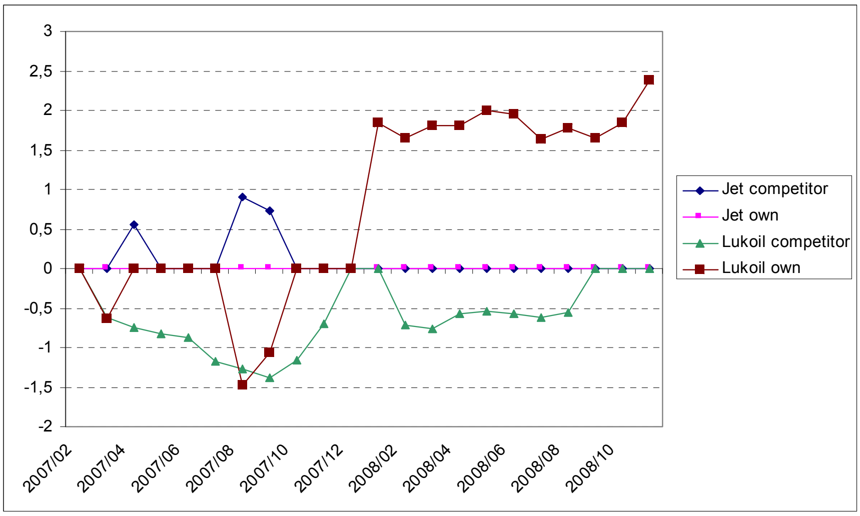
As a final sensitivity check we look at the margin differences between control and treatment groups throughout the full sample period. More precisely we estimate the monthly price differentials for the stations of the four merging firms and of their competitors with the specification of Equation (2) and we normalize the monthly price differentials to zero in the month of the clearance for the firms affected by the respective merger, (February 2007 for Lukoil/Jet and July 2007 for Agip/Esso). Graphs 7 and 8 show the evolution of the margin differentials, where we plot only parameter estimates that are significant at 5%.

These normalized control-treatment differences are informative in two respects. First, they offer a way to test for the validity of the assumptions behind the difference-in-differences estimator. As Imbens and Wooldridge (2009) point out, this normalized difference should be zero in the pre-merger periods if the difference-in-differences assumptions are valid. Looking at the graphs shows that on average this condition is satisfied for all merger effects but the Lukoil competitor effects. This suggests that the Lukoil competitor estimates should be interpreted carefully, since the estimator might be biased. We did not have significant estimates for this effect, however, therefore our core results are not affected by this finding.

Graph 7. Significant monthly differentials of treatment and control stations for Agip and Esso and their competitors



Graph 8. Significant monthly differentials of treatment and control stations for Lukoil and Jet and their competitors



Second, the normalized differences provide further insight on the merger date choice, because they basically show what the 'before' and 'after' averages are calculated from: the standard difference-in-differences estimator of any merger effect is the difference between the average of the respective monthly differentials in months before and after the effective merger date. The results show a visible difference emerging between the prices of treatment and control stations for both mergers from the end of 2007. In the case of the Agip/Esso merger, the price differentials are steadily increasing for the treatment group belonging to Esso own and competitor effects, which can be consistent with the gradual brand repositioning explanation we discussed earlier. In the case of stations used in estimating the Lukoil own effect, however, there is a discrete jump in the monthly price differentials between stations in the treatment and control group, which can indicate a sharp change in the firms' pricing policy.

The graphs show why some results are more sensitive to the choice of the effective merger date than others. For example, as the monthly differential is almost never significant in the case of Jet, the difference-in-differences estimator robustly shows zero Jet own and competitor effects to all effective merger dates; while the increase of Esso monthly differentials explains why we estimate larger Esso effects if the effective merger date lies further away from the clearance date.

3.10. Conclusion

This paper showed how to separate the ex post effects of simultaneous mergers on the prices of buyer and seller firms and their competitors. We exploit variation in the combination of affected firms' presence in distinct local markets to identify and estimate these effects by difference-in-differences methods. The separation of these effects enables us the testing of previous theoretical predictions of the merger literature explaining how the firms' different role in the merger may result in different price changes after the merger.

As an application, we used a sufficiently rich panel database of station-based prices to analyze two almost simultaneous mergers in the Hungarian retail gasoline market. We concluded that a positive but negligible price effect can be attributed to both mergers, but one merger resulted in higher prices for the buyer firm's stations only, while the other increased prices of seller's stations and of its competitors. We also checked whether these results were sensitive to the (unobservable) dates when the mergers effectively changed the firms pricing, and found that our qualitative implications emerge robustly.

Our method for separating the different price effects of mergers can be applied to any merger where there is some variation in the activities of the affected firms among distinct local markets (or in some cases, among distinct products). Therefore, given appropriate data, one could use this approach in the ex post evaluation of airline, hospital or supermarket mergers, which typically attract public and political attention. The method can also be modified to study research questions emerging from the specific needs of a policy case. In our application, for example, we could have easily studied how the mergers affected the pricing of stations owned by vertically integrated companies and individual stations differently by appropriately selecting the local markets identifying this effect.

In the future, we plan to complement our database with data on driving distances between stations. This feature will allow us to fine-tune the local market approach we have been working with, and check the robustness of our results in this respect. We can also add a further set of station characteristics to control for additional services like dining or car wash facilities, which could shed further light on the competition between leading brands offering a full range of services and discount stations supplying only gasoline.

References

[1] Ackerberg, D., L. Benkard, S. Berry and A. Pakes (2007). "Econometric Tools for Analyzing Market Outcomes," in: J.J. Heckman & E.E. Leamer (ed.), *Handbook of Econometrics*, edition 1, volume 6, chapter 63 Elsevier.

[2] Aguirregabiria, V. and P. Mira (2007). "Sequential Estimation of Dynamic Discrete Games," *Econometrica*, 75,1–53.

[3] Andreoni, J. (1989). "Giving with Impure Altruism: Applications to Charity and Ricardian Equivalence," *Journal of Political Economy*, 97(6), 1447-58.

[4] Andreoni, J. (1990). "Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving?," *Economic Journal*, 100(401), 464-77.

[5] Álvarez, L. J., E. Dhyne, M. Hoeberichts, C. Kwapil, H. Le Bihan, P. Lünnemann, F. Martins, R. Sabbatini, H. Stahl, P. Vermeulen and J. Vilmunen (2006). "Sticky Prices in the Euro Area: A Summary of New Micro-Evidence," *Journal of the European Economic Association*, 4(2-3), 575-584, 04-05.

[6] Armantier. O. and O. Richard (2008). "Domestic Airline Alliances and Consumer Welfare," *RAND Journal of Economics* 38, 875-904.

[7] Ashenfelter, O. and D. Hosken (2008). "The Effect of Mergers on Consumer Prices: Evidence From Five Selected Case Studies," NBER Working Paper No. 13859.

[8] Ashenfelter, O., Ashmore, D., Baker, J., Gleason, S. and Hosken D. (2006). "Empirical methods in merger analysis: Econometric analysis of pricing in FTC v. Staples," *International Journal of the Economics of Business* 13, 265–79.

[9] Asplund, M., R. Eriksson and R. Friberg (2000). " Price Adjustments by a Gasoline Retail Chain," *Scandinavian Journal of Economics* 102(1), 101-21.

[10] Bachmeier, L. J., J. M. Griffin (2003). "New Evidence on Asymmetric Gasoline Price Responses," *The Review of Economics and Statistics* 85(3), 772-776, 05.

[11] Bacon, R. W. (1991). "Rockets and feathers: the asymmetric speed of adjustment of UK retail gasoline prices to cost changes," *Energy Economics*, 13(3), 211-218.

[12] Bajari, P., H. Hong, J. Krainer and D. Nekipelov (2010). "Estimating Static Models of Strategic Interactions," *Journal of Business & Economic Statistics, American Statistical Association* 28(4), 469-482.

[13] Bajari, P., H. Hong and D. Nekipelov (2010). "Game Theory and Econometrics: A Survey of Some Recent Research." Econometric Society World Congress in Shanghai.

[14] Bernheim, B. D. (1994). "A theory of conformity" *The Journal of Political Economy* 102. No. 5. (Oct., 1994), pp.841-877.

[15] Bettendorf, L., S.A. van der Geest and M. Varkevisser (2003). "Price asymmetry in the Dutch retail gasoline market," *Energy Economics* 25(6), 669-689.

[16] Bjørner, T. B., L. G. Hansen and C. S. Russell (2004). "Environmental labeling and consumers' choice - an empirical analysis of the effect of the Nordic Swan," *Journal of Environmental Economics and Management* 47, 411-434.

[17] Blend, J. R. and E. O. van Ravenswaay (1999). "Measuring Consumer Demand for Ecolabeled Apples" *American Journal of Agricultural Economics* Vol. 81, No. 5, Proceedings Issue , pp. 1072-1077

[18] Bodner, R. and D. Prelec (2002). "Self-signaling in a neo-Calvinist model of everyday decision making," in *Psychology and Economics*, Vol. I. I. Brocas and J. Carillo (eds.), Oxford University Press,

[19] Borenstein, S., A. C. Cameron and R Gilbert (1997). "Do Gasoline Prices Respond Asymmetrically to Crude Oil Price Changes?," *The Quarterly Journal of Economics* 112(1), 305-39.

[20] Borenstein, S. and A. Shepard (2002). "Sticky Prices, Inventories, and Market Power in Wholesale Gasoline Markets," *RAND Journal of Economics* 33(1), 116-139.

[21] Brekke, K. A., S. Kverndokk and K. Nyborg (2003). "An economic model of moral motivation," *Journal of Public Economics* 87 1967–1983.

[22] Bresnahan, T. and P. Reiss (1991). "Entry and Competition in Concentrated Markets," *Journal of Political Economy* 99, 977-1009.

[23] Brock, W. A. and S. N. Durlauf (2003). Multinomial Choice with Social Interactions. NBER Working Paper No. T0288.

[24] Brock, W. A. and S. N. Durlauf (2001). "Discrete Choice with Social Interactions," *Review of Economic Studies*, 68(2), 235-60.

[25] Campbell, J, and H. Hopenhayn (2005). "Market Size Matters," *Journal of Industrial Economics* 53, 1-25.

[26] Choné, P. and L. Linnemer (2010). "A treatment effect method for merger analysis with an application to parking prices in Paris," working paper CREST.

[27] Dafny, L. (2009). "Estimation and identification of merger effects: An application to hospital mergers," *Journal of Law and Economics* 52, 523–550.

[28] Davis, P. and E. Garces (2009). *Quantitative Techniques for Competition and Antitrust Analysis*, Princeton University Press.

[29] Deltas, G. (2008). "Retail gasoline price dynamics and local market power," *Journal of Industrial Economics* 56(3), 613-628, 09.

[30] Deneckere, R. and C. Davidson (1985). "Incentives to Form Coalitions with Bertrand Competition," *RAND Journal of Economics* 16, 473-486

[31] Eckert, A. (2002). "Retail price cycles and response asymmetry," *Canadian Journal of Economics* 35(1), 52-77.

[32] Faber, R. P. (2009). "Asymmetric Price Responses of Gasoline Stations: Evidence for Heterogeneity of Retailers," Tinbergen Institute Discussion Papers 09-106/1, Tinbergen Institute.

[33] Farrell. J., P. Pautler and M. Vita. (2009). "Economics at the FTC: Retrospective Merger Analysis with a Focus on Hospitals," *Review of Industrial Organization* 35, 369–385.

[34] Fehr, E. and S. Gächter (2000). "Fairness and Retaliation: The Economics of Reciprocity" *The Journal of Economic Perspectives* 14, 3, 159-181.

[35] Fischbacher, U., S. Gächter and E. Fehr (2001). "Are People Conditionally Cooperative? Evidence from a Public Goods Experiment," *Economics Letters*, 71, 397-404.

[36] Fehr, E. and S. Gächter (2002). "Altruistic punishment in humans," *Nature* 415, 137-140.

[37] Frey, G. & M. Manera (2007). "Econometric Models Of Asymmetric Price Transmission," *Journal of Economic Surveys*, 21(2), 349-415, 04.

[38] Focarelli, D. and F. Panetta (2003). "Are Mergers Beneficial to Consumers? Evidence from the Market for Bank Deposits," *American Economic Review* 93, 1152-1172.

[39] Geweke, J. (2004). "Issues in the “Rockets and Feathers” Gasoline Price Literature" Report to Federal Trade Commission, March 16, 2004. http://www.ftc.gov/bc/gasconf/comments2/gewecke2.pdf

[40] Grasso, M. & M. Manera (2007). "Asymmetric error correction models for the oil-gasoline price relationship," *Energy Policy* 35(1), 156-177.

[41] Hastings, J. (2004). "Vertical Relationships and Competition in Retail Gasoline Markets: Empirical Evidence from Contract Changes in Southern California," *American Economic Review* 94, 317-328.

[42] Hortacsu. A. and C. Syverson (2007). "Cementing relationships: Vertical integration, foreclosure, productivity, and prices," *Journal of Political Economy* 115, 250–301.

[43] Hunter, G., G Leonard and S. Olley (2008). "Merger Retrospective Studies: A Review," *Antitrust* 23, 34-41.

[44] Imbens, G. and J. Wooldridge (2009). "Recent Developments in the Econometrics of Program Evaluation," *Journal of Economic Literature* 47, 5-86.

[45] Kim. E. and V. Singal (1993). "Mergers and market power: Evidence from the airline industry," *American Economic Review* 83, p. 549–69.

[46] Kotchen, M.J. (2006). "Green Markets and Private Provision of Public Goods," *Journal of Political Economy*, 114(4), 816-845.

[47] Lancaster, K., (1971). *Consumer Demand: A New Approach.* Columbia University Press. New York & London.

[48] LEAR (2006), Ex-post review of merger control decisions.

[49] Levy, D. and J. Reitzes (1992). "Anticompetitive Effects of Mergers in Markets with Localized Competition," *Journal of Law, Economics and Organization* 8, 427-440.

[50] Lewis, M. (2011*)*. “Asymmetric Price Adjustment and Consumer Search: An Examination of the Retail Gasoline Market”, *Journal of Economics and Management Strategy* 20(2), 409-449, 06.

[51] Lewis, M. & M. Noel (2011). "The Speed of Gasoline Price Response in Markets With and Without Edgeworth Cycles," The Review of Economics and Statistics, 93(2), 672-682, 01.

[52] Maskin, E. and J. Tirole (1988). "A theory of dynamic oligopoly II: Price competition, kinked demand curves and Edgeworth cycles," *Econometrica*, 56, 571–599.

[53] McFadden, D. (1974). "Conditional Logit Analysis of Qualitative Choice Behavior," [PDF file, 3.2M] in P. Zarembka (ed.), *Frontiers in Econometrics*, 105-142, Academic Press: New York, 1974.

[54] McFadden, D. (1981). "Econometric Models of Probabilistic Choice," in C.F. Manski and D. McFadden (eds.), *Structural Analysis of Discrete Data with Econometric Applications*, 198-272, MIT Press: Cambridge, MA, 1981.

[55] Meyer J. & S. Cramon-Taubadel (2004). "Asymmetric Price Transmission: A Survey," *Journal of Agricultural Economics* 55(3), 581-611.

[56] Nimon, W. & J. C. Beghin (2008). "Are Eco-Labels Valuable? Evidence from the Apparel Industry," Staff General Research Papers 12741, Iowa State University, Department of Economics.

[57] OECD, (2008). "Competition policy for vertical relations in gasoline retailing", DAF/COMP(2008)35

[58] Office of Fair Trading and Competition Commission (2005). "Ex post evaluation of mergers," OFT Economic Discussion Paper Series.

[59] Peltzman, S. (2000). "Prices Rise Faster than They Fall," *Journal of Political Economy* 108(3), 466-502.

[60] Pesaran, M. H. and R. Smith (1995). "Estimating long-run relationships from dynamic heterogeneous panels," *Journal of Econometrics*, 68(1), 79-113.

[61] Pesendorfer, M., and P. Schmidt-Dengler (2003). "Identifcation and Estimation of Dynamic Games," NBER working paper No. w9726.

[62] Prager, R. and T. Hannan (1998). "Do Substantial Horizontal Mergers Generate Significant Price Effects? Evidence from the Banking Industry," *Journal of Industrial Economics* 46, 433-452.

[63] Radchenko, S. (2005). "Oil price volatility and the asymmetric response of gasoline prices to oil price increases and decreases," *Energy Economics*, 27(5), 708-730.

[64] Reagan, P. B. and M. L. Weitzman (1982). "Asymmetries in price and quantity adjustments by the competitive firm," *Journal of Economic Theory*, 27(2), 410-420.

[65] Rege, M. (2004). "Social Norms and Private Provision of Public Goods," *Journal of Public Economic Theory* 6(1),65-77, 02.

[66] Roe, B., M. F. Teisl, A.Levy and M. Russell (2001). "US consumers' willingness to pay for green electricity," *Energy Policy* 29(11), 917-925.

[67] Rust, J. (1994). "Structural Estimation of Markov Decision Processes," in: *Handbook of Econometrics*, Volume IV, Edited by R.F. Engle and D.L. McFadden, Elsevier.

[68] Sapienza, P. (2002). "The Effects of Banking Mergers on Loan Contracts," *Journal of Finance* 57, 329–367.

[69] Simpson, J. and C. Taylor (2008). "Do Gasoline Mergers Affect Consumer Prices? The Marathon Ashland Petroleum and Ultramar Diamond Shamrock Transaction," *Journal of Law and Economics* 51, 135-152.

[70] Sugden, R.(1984). "Reciprocity: The Supply of Public Goods through Voluntary Contributions," *Economic Journal* 94(376), 772-87.

[71] Taylor, C. and D. Hosken, (2007). "The Economic Effects of the Marathon-Ashland Joint Venture: The Importance of Industry Supply Shocks and Vertical Market Structure," *Journal of Industrial Economics* 55, 419-451.

[72] Taylor, C., N. Kreisle and P. Zimmermann (2010). "Vertical Relationships and Competition in Retail Gasoline Markets: Comment," *American Economic Review* 100, 1269-1276.

[73] Verlinda, J. A. (2008). "Do Rockets Rise Faster and Feathers Fall Slower in an Atmosphere of Local Market Power? Evidence from the Retail Gasoline Market," *Journal of Industrial Economics* 56:3, 581–612.

[74] Vita, M. and S. Sacher (2001). "The Competitive Effects of Not-for-Profit Hospital Mergers: A Case Study," *Journal of Industrial Economics* 49, pp. 63-84.

[75] Vives, X. (1999). *Oligopoly Pricing*, MIT Press.

[76] Wang, X. H. and J. Zhao (2007). "Welfare reductions from small cost reductions in differentiated oligopoly," *International Journal of Industrial Organization* 25(1), 173-185.