Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN

Economia

Ciclo XXVII

Settore Concorsuale di afferenza: 13/A4

Settore Scientifico disciplinare: SECS-P01

Essays in Applied and Labor Economics

Presentata da: Alessandro Saia

Coordinatore Dottorato	**Relatori**
Matteo Cervellati	Andrea Ichino Matteo Cervellati

Esame finale anno 2015

Wishing for bad weather: demand shocks and labor supply of Bologna Pizza Vendors

Alessandro Saia*

Abstract

Recent influential papers analyze labor supply behavior of taxi drivers (Camerer *et al.*, 1997; and Crawford and Meng, 2011) and suggest that reference-dependence preferences have an important influence on drivers' labor-supply decisions. In the spirit of previous works I analyze the labor supply behavior of Bologna Pizza Delivery Vendors. Unlike previous papers, I am able to identify an exogenous and transitory change in labor demand. Using high frequency data on orders and rainfall as an exogenous demand shifter, I invariably find that reference-dependent preferences play no role in their labor' supply decisions and the behavior of pizza vendors is perfectly consistent with the predictions of the standard model of labor' supply. The difference in findings is due to differences in empirical approach rather than in the data used.

---

*Contact: alessandro.saia@unibo.it. I am particularly grateful to Andrea Ichino. I would like to thank Maria Bigoni, Matteo Cervellati, Efrem Castelnuovo, Juan J. Dolado, Irene Ferrari, Margherita Fort, Riccardo Ghidoni, Giovanni Mastrobuoni, Chiara Monfardini, Giacomo Pasini, Alfonso Rosolia, Giulio Zanella and seminar participants at Bologna, EUI, BOMOPAV 2014 Meeting and XIII Brucchi Luchino Workshop.

1 Introduction

There is a growing list of influential papers that documents a series of behavioral anomalies that question the basic assumption of standard economic models. Several studies provide evidence of the fact that individuals tend to overweight the present than the future (Laibson, 1997), they make decision using only a subset of the available information (Dellavigna and Pollet, 2009) and their utility is also function of the utility of other people (Charness and Rabin, 2002). To assess the extent to which the core assumptions of standard economic models are correct is of crucial importance for academics and policymakers alike. If relevant, incorporating behavioral insights in the design of public policies not only could improve the effectiveness of existing policies but could also offer new tools (Chetty, 2015).

The implications of behavioral biases may be particularly important in the context of labor economics. If behavioral factors play a large role in explaining labor supply decision then policies that neglect these anomalies, are suboptimal. The current design of income taxes (or unemployment insurance policies) is entirely based on the assumption of standard models of labor supply: labor supply should increase in response to transitory and positive changes in earnings opportunities. Is this assumption valid?

Until recently, empirical research aimed to investigate the relationship between workers' effort and earnings opportunities was conducted using individual panel data available for standard workers. However, standard workers face variations in wages that are permanent, hence correlated with significant income effects. Moreover, they usually operate in setting in which (at the intensive margin) it is not possible to set their own effort. Then, while the external validity of these results is beyond any doubt what it is doubtful is the identification of the underlying labor supply behavior. Due to the serious limitations in the available data, the effort exerted by scholars to study workers labor supply decisions was gradually diminishing. But in the last years the very large literature in economics studying the behavior of the underlying labor supply has regained a central position in the academic debate.

Recent influential papers analyze labor supply behavior of piece rate workers, such as taxi drivers (Camerer *et al.*, 1997; Farber, 2008; and Crawford and Meng, 2011) and bike messengers (Fehr and Goette, 2007). These specific groups of workers have proven to be the ideal candidates to investigate labor supply decision: they operate in settings in which they face transitory changes in their earnings opportunities and they are free to set their effort. However, despite the possibility to set their effort, these papers have provided evidence that is in sharp contrast with the prediction of the standard model but is consistent with alternative theories that explain labor supply decisions with reference-dependence preferences. Reference-dependence preferences (henceforth, *RDP*) models of labor supply behavior predict that workers labor supply depends not only on earnings opportunities as in standard neoclassical theory, but also on a target (*reference point*). In the case of taxi driver, the idea is that she decides to stop driving based on having reached the reference point (e.g. income goal) and not on whether future potential earnings are high or low.

Crawford and Meng (2011) is the first paper that, by investigating the daily labor supply of New York City taxi drivers, conceptualize drivers target as endogenous rational expectations. Building on Kőszegi and Rabin's (2006) theory of RDP, the authors estimate a model of daily labor supply as a survival model in which at the end of each trip drivers decide whether to continue or to stop driving. They derive a reduced-form model of the probability of stopping as linear function of cumulative hours and cumulative income. In order to investigate the relevance of reference-dependent behavior, they implement the probit model including two dummy variables indicating whether hours or income exceed the targets. Targets reflect the driver' rational expectations in terms of hours and income and represent the belief about outcomes that she might face. Since targets are not observed, they operationalize them as functions of endogenous variables. Given the lack of suitable demand-shifting instruments, they compute sample proxies for driver' rational expectation using driver specific sample average income and hours prior to the current shift, allowing them to vary across days of week in order to capture variation in demand throughout the week. In their paper, they provide evidence of deviations from the standard neoclassical model of worker behavior, suggesting that the behavioral component is an important part of the labor-supply decisions of taxi drivers. However, why effort exerted by taxi drivers varies through the day is no clear and Crawford and Meng (and none of the previous papers) were able to identify an exogenous demand shifter that could be used as instrument to provide credible estimates of the sign of the labor supply elasticity.

In the spirit of previous works I analyze the labor supply behavior of Bologna Pizza Delivery Vendors using high frequency data on orders placed over the period February 2010-July 2011. At the intensive margin (i.e. during a day), vendors, like taxi drivers, can decide how much effort they want to exert by accepting (or not accepting) orders. A key feature of this paper is that I am able to identify an exogenous and transitory shift in labor demand. Among other determinants, weather conditions are one of the key shifters of demand for pizza delivery services. Using Web search data I will provide evidence of the fact that rainfall are a statistically significant determinant of demand of pizza delivery. Then, unlike previous papers, I am able to instrument the effort of Pizza delivery vendors using exogenous and transitory change in weather conditions. Using changes in rainfall as sources of exogenous change in pizza delivery demand, I invariably find that vendor's behavior is consistent with the prediction of the standard model of labor supply. This result contrasts sharply with the results of previous studies and I show that this difference in findings is due to differences in analytical approach rather than in the data used.

The remainder of this paper is structured as follows. The next section describes the data utilized in this paper and describes the labor market for Pizza Delivery Vendors. Section 3 presents the empirical results on the relevance of adverse weather condition for pizza demand using Web searches data. In Section 4, I replicate the analysis of Crawford and Meng (2011) using Pizza Vendors' data and I obtain results that are consistent with those found for taxi drivers. In Section 5, I present the empirical results obtained exploiting the exogenous variation in rainfall and I show that the behavior of pizza delivery vendors is perfectly consistent with the predictions of the standard model of labor supply. Section 6 concludes.

2 Data and Background

The data used in this paper consists of high frequency data on orders carried out through an on-line takeaway pizza delivery service (henceforth, *the Website*). The Website acts as a web-based intermediary between privately owned restaurants and customers (it can be considered to be a modern variant of the *Yellows Pages*). The Website allows customers to select a pizza delivery restaurant, place the order and then receive the delivery at home. When a customer places an order, a message is sent to the delivery restaurant using a 2-way till receipt/order machine. Vendors receive the order through the machine box and then are able to accept or reject the order. If they accept it, a confirmation message is sent to the customer with an indication of the delivery time (set by the vendors). Payments are collected by the vendors on delivery and for each order the Website charges a flat commission rate.

Pizza delivery restaurants are located in Bologna and are privately owned family business. Unlike normal restaurant, their distinguishing feature is that they do not provide table service and the majority of customer orders are carried out by telephone or over the web. Therefore, I observe only a part of the total demand (i.e. the orders received through the Website) but it reasonable to assume that the two channels (telephone and web) are positively related.

Unlike previous works, this paper investigates the labor supply behavior of a group of workers. I consider each single restaurant as a group of workers and I will refer to all workers of a single restaurant as vendors. Moreover, I assume that the group (in terms of size and composition) is constant through all the period. Given labor market characteristics (i.e. restaurants are family business) it is reasonable to assume this to be true. In the following section I replicate the analysis of Crawford and Meng (2011) using pizza delivery data. Obtaining similar finding is of crucial importance because it implies that despite the intrinsic differences between taxi drivers and pizza vendors (single vs group of workers, settings and cities) any difference in the results is due to the different econometric approach rather than to differences between the two settings.

2.1 Data on Pizza Sales

The Website provided me all orders placed over the period February 2010-July 2011 for 59 pizza delivery restaurants. As explained in the previous section vendors receive the order through a machine box and then are able to accept or reject it. Rejected orders (as consumers that weren't able to find a taxi) are crucial to provide a credible estimate of intertemporal substitution of labor supply (Frisch elasticity). Unfortunately, as in previous work, in this paper I rely only on the equilibrium outcomes, due to the fact that rejected orders weren't recorded.

From February 2010 to July 2011 a total of 147.090 orders were observed. Figure 1 plots the location of each pizza delivery restaurants across the city of Bologna (red dots) and the spatial distribution of orders observed in the sample (blue dots). For each order I have detailed information on the date and time the order was placed, the total amount of the delivery, the distance between the restaurant and the delivery place and the expected delivery time, that is communicated to customers when the order is accepted by the vendors.

[Figure 1]

Daily average earnings across the whole sample are equal to 68.15 and the average number of orders accepted in a day is around 8. Not surprisingly, daily earnings opportunities vary across days of the week and across months. Earnings opportunities are higher during winter and during the weekend (for example, daily earnings and daily orders observed during Sundays of January are almost twice that of the overall sample). Figure 2 contains kernel density of the distribution of orders over the day. The number of order received varies greatly throughout the day: orders are concentrated around lunch breaks (generally 12:30–3:00 pm) and dinner hours (generally 6:30–9:00 pm). This pattern highlights the crucial role of meal times in demand of pizza delivery services.

[Figure 2]

2.2 Weather Data

Data on weather conditions over the entire period in the city of Bologna are provided by the Regional Agency for the Protection of the Environment (ARPA) of Emilia Romagna. The data on hourly precipitation is drawn from a meteorological station located in the city centre of Bologna (*Bologna Urbana*). Note that the city of Bologna, with a population of about 384.000 and a total area of about 140.7 km^2, is relatively small[1].

[Figure 3]

Figure 3 plots the number of days per month with measurable precipitation. From February 2010 to July 2011 the number of rainy days (defined as days with 1 mm or more of rain) is about 206.

3 The Demand for Pizza Delivery in Bologna

Demand for pizza delivery fluctuates on a daily basis due to demand shocks caused by day-of-the-week effects, sport events, holidays, etc. Among others, weather conditions are one of the key shifters of demand for pizza delivery services. Workers of the food delivery industry recognize that rainfall are strongly correlated with positive short-run fluctuations in demand for pizza delivery services. For example, in 2012, while announcing a rise in firm' profits the CEO of Domino's Pizza[2], Lance Batchelor said " *When the weather turns nasty we always see sales turn upwards*".

In this section I use Web search volume to provide evidence of the positive effect of adverse weather condition on pizza demand. Recently several papers have demonstrated that Web search data correlates in near real time with individual activity and that search counts are predictive of economic outcomes, including unemployment levels (Ettredge *et al.*, 2005), home sales (Wu and Brynjolfsson, 2013) and consumers activities such as movie attendance and music sales (Goel *et al.*, 2010). In the same spirit, I use daily Web search volume to derive information about the elasticity of pizza demand to rain conditions. Pizza related search queries are provided by Google Trends. Google Trends provides me an index[3] that indicates the daily volume of queries entered into Google that

[1] To give just one example, in New York there are over 8.4 million residents in 1,214 km^2 and the distance between Central Park (where it is located the weather station used in Crawford and Meng and Farber) and Kennedy Airport (a pick-up location associated with the highest earnings) is of about 21.5 km

[2] Domino's Pizza is an American restaurant chain and international franchise pizza delivery corporation with 145.000 employees and more than 1 billion of USD of revenue.

[3] Google Trends data does not report the raw level of queries for a given search term, rather, it reports a query index, that indicates the volume of searches for the specific search term divided by the total number of queries in a given geographic location at a certain point in time.

includes the Italian words for "pizza restaurant Bologna" in the area of Bologna.

Figure 4 plots the Google Trends daily index of the search "pizza restaurant Bologna" from February 2010 to July 2011. The fluctuation of searches is consistent with the behavior of pizza demand: searches peak during cold seasons and follow a seasonal trend that is compatible with pizza consumption.

[Figure 4]

I investigate to which extent adverse weather conditions affects daily Web search volume using the following model:

$$log(Searches_t) = \beta + \beta_1 Rainy\ Day_t + \epsilon_t \tag{1}$$

where $Searches_t$ represents the index of searches in day t and *Rainy Day* is a dummy variable that takes value equal to 1 if in the corresponding day it has rained. Column (1) of Table (1) reports the estimation results from the specification with day of the week, month and year dummies. As expected the coefficient of interest is positive and statistically significant: adverse weather conditions during the day increase roughly by 6% search queries related with pizza consumption.

[Table 1]

Not surprisingly, results are robust when investigating the effect of precipitation at the intensive margin. Results in Columns 2-4 of Table (1) present the estimates using as alternative variable of interest the total amount of rainfall (in millimeters), the average precipitation in the day (in millimeters) and the number of rainy hours in the day. For example, three hours of rain are associated with an increase of around 2% in the daily volume of pizza related queries entered into Google.

The randomness in weather events is function of precipitation intensity and time of the day. Since meal times are among the main determinants of pizza delivery I investigate the potential heterogeneous impact of rainfall on pizza demand during a day. Using the distribution of the orders in my sample I have derived weights for each hour. Weights reflect pizza delivery demand intensity throughout the day. Hourly precipitations observed in the day are weighted using the fraction of total orders observed in the corresponding hour over the entire period. The underlying idea is to investigate if the impact of adverse weather conditions varies depending on the time of the day.

Columns 1-3 of Table 2 present the estimates of the model where the variables indicating adverse weather conditions are defined as the sum of the product between millimeters of rain and the corresponding hour weights, as the sum of the interaction of rainy hours and weights and as the average of the product between millimeter of rain in the day and the corresponding hour weights, respectively.

[Table 2]

Taken as a whole, these results demonstrate that rainfall are a statistically significant determinant of demand of pizza delivery. The positive significant effect reported by the estimated coefficients indicating adverse weather condition in the day, provide further evidence of the positive effect of rain on the demand of delivery food, as discussed before.

4 Estimation of Labor Supply Models

4.1 Estimation using Crawford and Meng' Targets

In the next section I replicate the analysis of Crawford and Meng (2011) using pizza sales data. Obtaining similar finding is of crucial importance because it implies that despite the intrinsic differences between taxi drivers and pizza vendors (e.g. single vs group of workers, settings and cities) any difference in the results is due to the different econometric approach rather than to differences between the two settings.

I depart from their analysis modelling labor' decision of pizza vendors solely as a function of their effort[4]. Pizza vendors are piece-rate workers and their labor supply decision is about how much effort they want to exert by accepting orders. Then, as a proxy of vendors' effort, I use the number of order accepted in the day[5]

Using pizza-delivery data, I estimate a model of labor supply where the decision of stopping is function of cumulative orders accepted in the day and a dummy variable indicating whether pizza produced in the day exceeds the target. Following Crawford and Meng (2011) I compute the target using the average number of previous orders accepted. Despite the presence of endogenous regressors the empirical strategy proposed by Crawford and Meng (2011) has the advantage to isolate the neoclassical component of labor supply from the reference-dependent effect. While the estimated coefficients of cumulative orders provides evidence of the part of labor-supply behavior related with standard preferences, the coefficient of the dummy indicates the variation in the continuation probability when target is reached. If positive and statistically significant, the coefficient of the dummy will provide evidence of target behavior by pizza vendors

[4]However, it is important to note that the fundamental results of this section are the qualitatively the same when I adopt a model of labor' supply with two targets, defined as specific sample average income and hours prior to the current day, as the one used by Crawford and Meng (2011).

[5]In principle I could use the number of pizza produced or the total earnings in the day as alternative measures of effort. The results do not change substantively.

Table 3 contains estimates of reduced form model of stopping probability allowing jump at the target[6]. In the left-hand panel targets are computed by sample averages, vendors/month by vendors/month, prior to the day in question. The model in the first column contains only cumulative orders accepted and the dummy indicate whether number of orders is above the target. In the second column the model includes controls for vendors-day-of-the-week' heterogeneity, month, year, hour of the day and weather condition. In both models target coefficients suggest the presence of relevant deviation from the standard model of labor supply and suggest that vendors are more likely to stop working when they hit their target. At the same time, the level of orders accepted has a significant and negative effect providing evidence of the presence of a neoclassical component in the labor supply decision of pizza vendors. The right-hand panel reports estimates obtained computing target by vendors, month and day of the week. Overall the estimates confirm previous results: when the demand is higher than expected (i.e. vendors hit their targets), vendors are more likely to end working.

[Table 3]

Despite the differences in the two settings the results presented in this section are substantially consistent with those found by Crawford and Meng (2011). Pizza vendors look like taxi drivers in terms of inter-temporal labor supply elasticity. As for taxi drivers, being above the target increases the probability of quit and reference-dependent preferences seem to play a key role in labor supply behavior of pizza vendors. The similarity of findings is crucial because it implies that any differences in future results obtained with alternative empirical strategy are due to differences in the analytic approach rather than to differences in the two settings.

Not surprisingly, the coefficient of the dummy variable indicating whether during the corresponding hour it was raining is negative and statistically significant. That is to say that when earnings opportunities are temporarily high vendors behavior is consistent with a standard model of labor supply and workers effort is positively related with transitory variation in demand[7]

It should be noticed that a prediction of the role of relevance of reference dependent preferences in labor supply decision is that workers should respond differently to demand shock if they are above the target[8]. Differently from Crawford and Meng (2011) I can inves-

[6]In the main text are presented results from unweighted regression, where sampling variation is ignored. In order to take into account sampling variation, I produce (as in Crawford and Meng, 2011) weighted estimates, where weights are equal to the number of observation used to compute the target. The corresponding results are reported in Appendix 3. Results obtained with unweighted and weighted regression are essentially the same.

[7]Similar results are obtained from when rain defined in millimeters.

[8]In the Appendix 2 I develop two alternative models of labor supply for piece-rate workers that operate in a market with menu costs to derive theoretical predictions.

tigate this theoretical prediction through the interaction between the dummy indicating whether orders exceed the target and the variable indicating if in the corresponding hour it was raining. If reference dependence plays an important role in determining vendor's labor supply decisions, the coefficient of the interaction term is expected to be positive and statistically significant: if vendors have reached their target they should respond less to an increase of pizza delivery demand.

[Table 4]

The corresponding results are reported in Table 4 and do not provide any support for the relevance of reference dependent preferences in vendors' labor supply decisions. The coefficients of the interaction terms are not statistically significant at conventional levels. When target is computed as vendors/month average orders accepted (Column 1) the point estimate of the coefficient is negative. When target is defined as vendors/day/month averages (Column 2) the coefficient is actually positive, as theory to predict, but far from being significant.

Takes as a whole the results provided in this section indicate that *i*) pizza vendors look like taxi drivers and target behavior seem to play a key role in their labor supply decisions *ii*) adverse weather conditions are positively related with workers effort and *iii*) the role of target is unclear when a truly exogenous demand shifter is used.

4.2 Estimation using Rainfall Variation

Unlike Crawford and Meng I can investigate the relevance of the target component using demand shocks, defined in term of rainfall variation, as an instrument for earnings opportunities. In this section, I exploit the change in weather conditions to isolate exogenous variations in vendors' earnings opportunities. Using precipitation in the day I identify two instruments. I compute *Cumulative Rainy Hours* as the number of rainy hours observed up to that point in the day and a target variable labeled *Target Rainy Hours*, defined as vendors/month and vendors/month/day of week sample average observed rain prior to the current day as target. The first variable isolates the neoclassical component of labor supply, while the second variable is a proxy for the vendors' rational point expectations of a day's pizza demand[9].

For each pizza vendors, cumulative precipitation are computed using a given hour as starting point. The starting hour corresponds to the earliest hour at which an order was observed over the entire sample. Thus, while the starting hours vary among vendors,

[9]Due to the non-trivial solution of a standard two stage regression in a probit model with multiple endogenous variables I focus solely on the reduced form estimates. However, the information underlying the first-stage regressions is provided in the Appendix 4.

it is assumed to be constant over the period. An alternative approach is to consider as starting point the time at which the first order was observed in the day. This approach, implicitly adopted by Crawford and Meng, allows starting hours to vary among vendors and across days but relies on the strong assumption that the waiting time for the first order is always equals to zero. This approach seems inconsistent with daily activity of pizza vendors that, as for taxi drivers' one, requires a series of preliminary steps (e.g. preparing pizza requires, for example, a certain stable oven temperature)[10].

[Table 5]

Table 5 presents the estimates of the reduced form model where the probability that vendors end their work day is function of the number of rainy hours observed in the day and a dummy variable indicating whether precipitation in the day hit the estimated target. In the left-hand panel, vendors target is computed using vendors/month sample averages up to but not including the day in question. When defined as the sum of previous precipitation the coefficient of cumulative rain is negative and statistically significant. This results is similar to the evidence obtained in Section 4.1 and suggests the presence of a neoclassical component in the labor supply decision of pizza vendors.However, the estimated coefficient of dummy variables indicating whether precipitation in the day exceeds the target is close to 0 and far from being statistically significant. Similar findings are reported in the right-hand panel, where exogenous target is computed using vendors/month and day-of-week sample averages up to but not including the day in question. These estimates are inconsistent with previous results and provide no support for the relevance of reference-dependence preferences in the labor-supply story of pizza vendors.

In Table 6 are reported the marginal probability effects obtained using a weighted measure of cumulative precipitation as proxy of pizza demand for the day. Results obtained using Web search data highlighted that the impact of adverse weather conditions varies depending on the time of the day, whereas the effects are greater during meal times. In order to compute cumulative weighted precipitation, previous precipitation in the day are weighted using the fraction of total orders observed in the corresponding hour over the entire period[11]. Overall, estimates reported in Table 6 are consistent with the ones presented in the previous Table 5. The variables indicating exogenous changes in vendors' earnings opportunities are statistically significant, whereas target coefficients are not.

[Table 6]

Taken as a whole, these results are in contrast with the estimates of previous section.

[10] It is worth to notice when the alternative measure is adopted, the results are also qualitatively similar to those presented in this section.

[11] In order to avoid endogeneity problems, vendors' weight are computed using the distribution of total orders of other vendors

When truly exogenous demand shifters are used, the behavior of pizza vendors is perfectly consistent with the predictions of the standard model of labor supply and reference-dependent preferences play no role in vendors' labor supply decisions.

5 Conclusion

In this paper I investigate the role of behavioral anomalies in the context of labor supply decisions. Using high frequency data on pizza sales I study the labor supply behavior of Bologna Pizza Delivery Vendors. Building on Crawford and Meng's (2011) analysis, I show that when targets are defined as average sample realizations of previous labor supply decisions I find evidence of reference-dependence behavior. Similarly to Crawford and Meng (2011), the probability of stopping work is higher when vendors hit their target. However, in contrast to previous works, I am able to identify an exogenous and transitory demand shifter. Unlike Crawford and Meng (2011), I can investigate the relevance of the target behavior using precipitation data as instrument for earnings opportunities. When truly exogenous demand shifters are used the target component plays no role in vendors's labor supply decisions and the observed behavior is consistent with the prediction of the standard model of labor supply.

References

Camerer, C., Babcock, L., Loewenstein, G., and Thaler, R. (1997). Labor supply of new york city cabdrivers: One day at a time. *The Quarterly Journal of Economics*, 112(2):407–441.

Chetty, R. (2015). Behavioral economics and public policy: A pragmatic perspective. *NBER Working Paper Series*, No. 20928.

Crawford, V. P. and Meng, J. (2011). New york city cab drivers' labor supply revisited: Reference-dependent preferences with rationalexpectations targets for hours and income. *The American Economic Review*, 101(5):1912–1932.

Ettredge, M., Gerdes, J., and Karuga, G. (2005). Using web-based search data to predict macroeconomic statistics. *Communications of the ACM*, 48(11):87–92.

Farber, H. S. (2005). Is tomorrow another day? the labor supply of new york city cabdrivers. *Journal of Political Economy*, 113(1):46–82.

Farber, H. S. (2008). Reference-dependent preferences and labor supply: The case of new york city taxi drivers. *The American Economic Review*, 98(3):1069–1082.

Farber, H. S. (2014). Why you can't find a taxi in the rain and other labor supply lessons from cab drivers. *NBER Working Paper Series*, No. 20604.

Fehr, E. and Goette, L. (2007). Do workers work more if wages are high? evidence from a randomized field experiment. *The American Economic Review*, 97(1):298–317.

Goel, S., Hofman, J. M., Lahaie, S., Pennock, D. M., and Watts, D. J. (2010). Predicting consumer behavior with web search. *Proceedings of the National Academy of Sciences*, 107(41):17486–17490.

Kőszegi, B. and Rabin, M. (2006). A model of reference-dependent preferences. *The Quarterly Journal of Economics*, 121(4):1133–1165.

Wu, L. and Brynjolfsson, E. (2013). The future of prediction: How google searches foreshadow housing prices and sales. *NBER Chapters*.

Tables and Figures

Figure 1: Pizza Delivery Restaurants and Orders Across the City of Bologna.

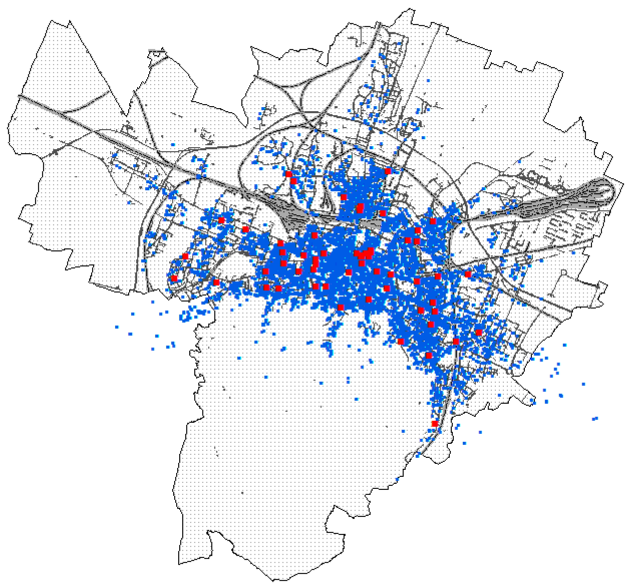


Figure 2: Distribution of Orders over the Day.

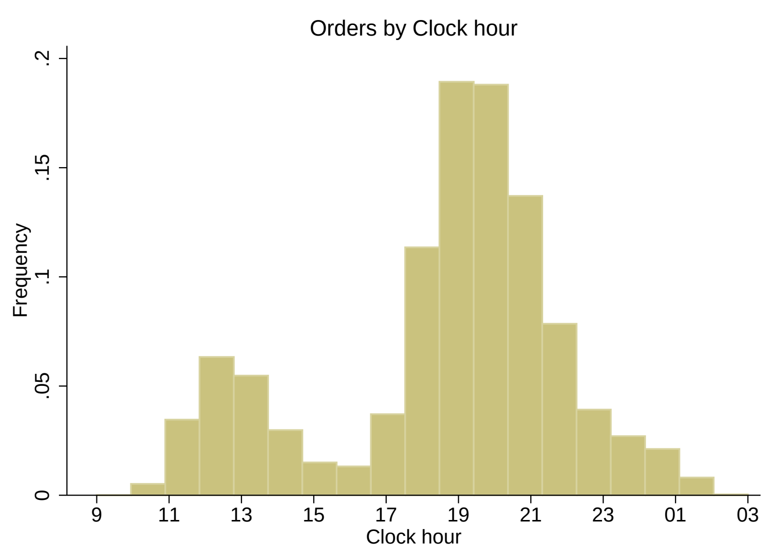


Figure 3: Number of Rainy Days per Month.

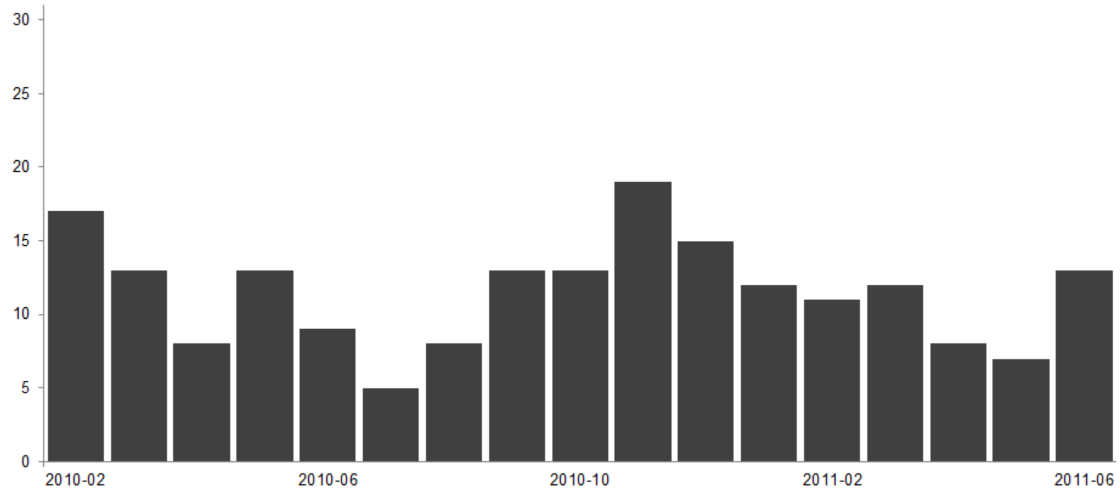


Figure 4: Daily Search Volume for the Words "Pizza Restaurant Bologna"

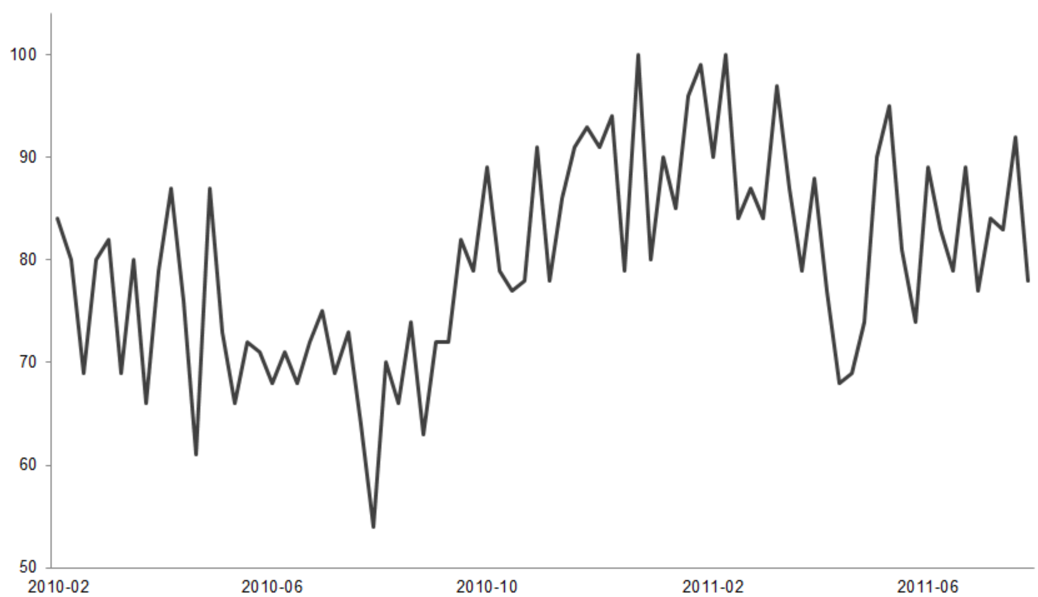


Table 1: Estimates of Google Searches Elasticity - Unweighted Precipitation

	(log) Searches	(log) Searches	(log) Searches	(log) Searches
Rainy Day	0.0635**			
	(0.026)			
Total Rainfall in the Day (in mm)		0.0069***		
		(0.002)		
Average Rainfall in the Day (in mm)			0.1671***	
			(0.055)	
Number of Rainy Hours				0.0072***
				(0.002)
Controls	Yes	Yes	Yes	Yes
Observations	546	546	546	546
R-squared	0.7030	0.7046	0.7046	0.7030

Notes: *** Significant at the 1% level, ** at the 5% level, * at the 10%. Robust standard error are in parentheses. Estimates based on data on daily search Google Index provided by Google Trends. I used the GI of the Italian word for "*pizza restaurant Bologna*" for the period February 2010 - July 2011. All columns include dummies for day of the week, month, year.

Table 2: Estimates of Google Searches Elasticity - Weighted Precipitation

	(log) Searches	(log) Searches	(log) Searches
Weighted Total Rainfall (in mm)	0.1134**		
	(0.030)		
Weighted Average Rainfall (in mm)		2.7224**	
		(0.724)	
Weighted Number of Rainy Hours			0.1463***
			(0.054)
Controls	Yes	Yes	Yes
Observations	546	546	546
R-squared	0.7040	0.7040	0.7025

Notes: *** Significant at the 1% level, ** at the 5% level, * at the 10%. Robust standard error are in parentheses. Estimates based on data on daily search Google Index provided by Google Trends. I used the GI of the Italian word for "*pizza restaurant Bologna*" for the period February 2010 - July 2011. All columns include dummies for day of the week, month, year.

Table 3: Marginal Effects on the Probability of Stopping - Crawford and Meng' Targets

	Using vendors and month specific sample average orders accepted prior to the current day as target		Using vendors, month and day-of-week specific sample average orders accepted prior to the current day as target	
	(1)	(2)	(3)	(4)
Cum. Orders > Target Orders	0.314***	0.033***	0.287***	0.023***
	(0.003)	(0.002)	(0.004)	(0.002)
Cumulative Orders	-0.009***	-0.004***	-0.007***	-0.003***
	(0.000)	(0.000)	(0.000)	(0.000)
Rainy Hour		-0.011***		-0.012***
		(0.003)		(0.003)
Controls	No	Yes	No	Yes
Log likelihood	-52,137.351	-34,539.514	-43,224.88	-28,521.724
Pseudo R^2	0.0751	0.3872	0.0704	0.3866
Observation	143,443	143,440	120,210	120,209

Notes: *** Significant at the 1% level, ** at the 5% level, * at the 10%. Standard error clustered by vendors and day are assumed. Columns (2) and (4) include dummies for vendors, day of the week, vendors and day of the week, month, year and hour of the day. The variable *Rainy Hour* takes value equal to 1 if in the corresponding hour it has rained.

Table 4: Heterogeneous Responses to Demand Shocks - Crawford and Meng' Targets

	Using vendors and month specific sample average orders accepted prior to the current day as target	Using vendors, month and day-of-week specific sample average orders accepted prior to the current day as target
	(1)	(2)
Above Target	0.220***	0.164***
	(0.013)	(0.014)
Cumulative Orders	-0.027***	-0.022***
	(0.001)	(0.001)
Rainy Hour	-0.075**	-0.099***
	(0.030)	(0.032)
Above Target * Rainy Hour	-0.004	0.024
	(0.039)	(0.043)
Controls	Yes	Yes
Log likelihood	-34,539.509	-28,521.592
Pseudo R^2	0.3872	0.3866
Observation	143,440	120,209

Notes: *** Significant at the 1% level, ** at the 5% level, * at the 10%. Standard error clustered by vendors and day are assumed. Columns (2) and (4) include dummies for vendors, day of the week, vendors and day of the week, month, year and hour of the day. The variable *Rainy Hour* takes value equal to 1 if in the corresponding hour it has rained.

Table 5: Marginal Effects on the Probability of Stopping - Target using Rainfall

	Using vendors and month specific sample average precipitation prior to the current day as target (1)	Using vendors, month and day-of-week specific sample average precipitation prior to the current day as target (2)
Cumulative Rainy Hours > Target Rainy Hours	0.0009	0.0047
	(0.002)	(0.003)
Cumulative Rainy Hours	-0.0173***	-0.0205***
	(0.004)	(0.005)
Controls	Yes	Yes
Log likelihood	-34,733.543	-28,630.311
Pseudo R^2	0.3838	0.3842
Observation	143,440	120,209

Notes: *** Significant at the 1% level, ** at the 5% level, * at the 10%. Standard error clustered by vendors and day are assumed. Columns include dummies for vendors, day of the week, vendors and day of the week, month, year and hour of the day.

Table 6: Marginal Effects on the Probability of Stopping - Target using Weighted Rainfall

	Using vendors and month specific sample average precipitation prior to the current day as target (1)	Using vendors, month and day-of-week specific sample average precipitation prior to the current day as target (2)
Cumulative Weighted Rainy Hours > Target Weighted Rainy Hours	-0.0015	0.0033
	(0.002)	(0.002)
Cumulative Weighted Rainy Hours	-0.1522***	-0.2505***
	(0.062)	(0.064)
Controls	Yes	Yes
Log likelihood	-34,737.158	-28,631.002
Pseudo R^2	0.3837	0.3842
Observation	143,440	120,209

Notes: *** Significant at the 1% level, ** at the 5% level, * at the 10%. Standard error clustered by vendors and day are assumed. Columns include dummies for vendors, day of the week, vendors and day of the week, month, year and hour of the day.

Appendix A1: Demand shocks and labor supply

Adverse weather conditions significantly increase pizza delivery demand. By exploiting the exogenous increase in food' delivery demand I am able to provide evidence of the sign of labor supply elasticity using the following model:

$$log(Orders_{it}) = \beta + \beta_1 Rainy\ Hour_t + \epsilon_i t \tag{2}$$

where dependent variable is the (log) number of orders accepted by vendors i in the hour t[12] and the variable *Rainy Hour* takes value equal to one if in the corresponding hour interval were observed adverse weather conditions. Table 7 reports the estimates obtained using two alternative measure of bad weather. The first column reports the variable of interest takes value equals to 1 if in the hour period were observed more than 0.00 cm of rain. The coefficient of interest is always positive and statistically significant. Pizza delivery demand is positively influenced by adverse weather conditions and, consequently, vendors react by exerting more effort. Similar results are obtained when the variable of interest is defined as the total amount of rainfall (in millimeters) in the hour interval (Column 2).

Taken as a whole, the estimates of Table 7 provide evidence of the positive relationship between labor demand and labor supply of food delivery vendors: vendors work more when earnings opportunities are higher. It should be noted, however, that these results, while providing evidence of the positive sign of the labor supply elasticity, are not very informative. Positive elasticity of labor supply is not inconsistent with the presence of target behavior. Vendors will not respond to a positive change in labor demand if, and only if, labor supply decision is exclusively driven by reference dependence preferences[13].

Table 7: Vendors' effort and demand shocks

Rainy Hour	0.045***	
	(0.007)	
Rain (in mm)		0.012***
		(0.003)
Controls	Yes	Yes
R^2	0.2990	0.2988
Observation	75,948	75,948

Notes: *** Significant at the 1% level, ** at the 5% level, * at the 10%. Standard error clustered by vendors and day are assumed. Column include dummies for vendors, day of the week, vendors and day of the week, month, year and hour of the day.

[12] Only hours with at least 1 order were considered.

[13] In a model of labor supply decisions where total utility is defined as a weighted average of consumption utility and gain-loss utility workers will not respond to demand shock only if the weight of the gain-loss utility's weight is equal to one. See Appendix A2.

Appendix A2: RDP and heterogenous responses to demand shocks

In the following sections I develop two models of labor supply for piece-rate workers that operate in a market with menu costs. Reference-dependence preferences model predicts that workers should respond differently to demand shock if they are above the target. Evidence of the reference dependent preferences in labor supply decision can be obtained by interacting the dummy variable indicating whether effort exceeds the target with the rainfall variable. The results reported in Table 4 do not provide such support.

A standard model of labor supply

Pizza vendors operate in a monopolistically competitive framework. The market is characterized by the following demand function:

$$q = (A - p) \tag{3}$$

where p and q represent price and quantity of pizza and $A>0$.
The vendor's utility is

$$U = p \cdot q - c(q) \tag{4}$$

where p and q represent price and quantity of pizza produced and $c(q) = \theta q$ represents linear effort cost.

Using (1) and (2) the vendor's maximization problem can be written as

$$\max_{q} U = (A - q) \cdot q - \theta q \tag{5}$$

The first order condition of (3) w.r.t. q is

$$\frac{\partial U}{\partial q} = A - 2 \cdot q - \theta = 0 \tag{6}$$

utility maximization yields

$$q^{\star} = \left(\frac{A - \theta}{2} \right) \tag{7}$$

and a price of

$$p^{\star} = \left(\frac{A + \theta}{2} \right) \tag{8}$$

Consider a demand shock ($A_R>A$) and suppose that the food vendor is able to change the price but only at a menu cost of z (i.e. the cost of literally printing new menus).
In the presence of menu cost the food vendor might find it optimal to **not** adjust its price

to the level

$$p_R^\star = \left(\frac{A_R + \theta}{2}\right) \tag{9}$$

rather, it might choose to keep its price at $p^\star$ and adjust its output to[14]

$$q_R^\star = \left(\frac{A_R - \theta}{2}\right) \tag{10}$$

In a market with the presence of menu costs a change in aggregate demand affects quantity but not prices.

A model with reference-dependent preferences

Following Kőszegi and Rabin (2006) I write the vendor's total utility $V(q|q^T)$, as a weighted average of standard utility $U(q)$ and gain-loss utility $R(q|q^T)$, with weights $1-\eta$ and η, as follows:

$$V(q|q^T) = (1-\eta)U(q) + \eta R(q|q^T) \tag{11}$$

where the gain-loss utility

$$R(q|q^T) = \mu(q - q^T) \tag{12}$$

where q^T is the effort target value[15],and $\mu(\cdot)$ is a "universal gain-loss function" that satisfies the following properties:

- $\mu(x)$ is strictly increasing
- $\mu''(x) \geq 0$ for $x<0$, and $\mu''(x) \leq 0$ for $x>0$

The vendor's maximization problem is

$$\max_q V(q|q^T) = (1-\eta)\left((A-q)q - \theta q\right) + \eta\left(\mu(q - q^T)\right) \tag{13}$$

The first order condition of (11) w.r.t. q is

$$\frac{\partial V}{\partial q} = (1-\eta)(A - 2q - \theta) + \eta\left(\mu'_q(q - q^T)\right) = 0 \tag{14}$$

[14] I observe such behaviour whenever $z > \Delta\tilde{U}$ where $\Delta\tilde{U}$ ($\Delta\tilde{U}$=$U_R - \tilde{U_R} = p_R \cdot q_R - C(q_R) - (p \cdot \tilde{q}_R - C(q_R))$.

[15] Following Kőszegi and Rabin (2006), I model target as set equals to the agents' rational expectation. Building on Crawford and Meng (2001), I treat them as point expectations rather than distributions $q^T = E(q) = \left(\frac{E(A)-\theta}{2}\right)$

I can identify the relation between A and q through [16]

$$\frac{\partial q}{\partial A} = -\frac{\frac{\partial \lambda}{\partial A}}{\frac{\partial \lambda}{\partial q}} = -\frac{1-\eta}{-(1-\eta)2 + \eta\left(\mu''_{qq}(q-q^T)\right)} \tag{15}$$

if $\eta = 1$ (i.e. the standard utility weight is equal to 0):

$$\frac{\partial q}{\partial A} = 0 \tag{16}$$

if $(q - q^T) < 0$ (i.e. the vendor is below his target):

$$\frac{\partial q}{\partial A}^{-} = -\frac{\frac{\partial \lambda}{\partial A}}{\frac{\partial \lambda}{\partial q}} = -\frac{1-\eta}{-(1-\eta)2 - \eta\left(\mu''_{qq}(\cdot)\right)} \tag{17}$$

if $(q - q^T) > 0$ (i.e. the vendor is above his target):

$$\frac{\partial q}{\partial A}^{+} = -\frac{\frac{\partial \lambda}{\partial A}}{\frac{\partial \lambda}{\partial q}} = -\frac{1-\eta}{-(1-\eta)2 + \eta\left(\mu''_{qq}(\cdot)\right)} \tag{18}$$

and

$$\frac{\partial q}{\partial A}^{+} < \frac{\partial q}{\partial A}^{-} \tag{19}$$

[16] We indicates λ equation (12)

Appendix A3: Weighted Regressions

Table 8: Marginal Effects on the Probability of Stopping - Crawford and Meng' Targets

	Using vendors and month specific sample average orders accepted prior to the current day as target		Using vendors, month and day-of-week specific sample average orders accepted prior to the current day as target	
	(1)	(2)	(3)	(4)
Cum. Orders > Target Orders	0.298***	0.025***	0.282***	0.018***
	(0.004)	(0.002)	(0.004)	(0.002)
Cumulative Orders	-0.008***	-0.003***	-0.007***	-0.002***
	(0.000)	(0.000)	(0.000)	(0.000)
Rainy Hour		-0.013***		-0.014***
		(0.003)		(0.003)
Controls	No	Yes	No	Yes
Log likelihood	-1,044,271.4	-692,525.65	-129,922.3	-86,127.294
Pseudo R^2	0.0737	0.3857	0.0721	0.3849
Observation	3,005,049	3,005,038	373,309	373,308

Notes: *** Significant at the 1% level, ** at the 5% level, * at the 10%. Standard error clustered by vendors and day are assumed. Columns (2) and (4) include dummies for vendors, day of the week, vendors and day of the week, month, year and hour of the day. The variable *Rainy Hour* takes value equal to 1 if in the corresponding hour it has rained.

Table 9: Heterogeneous Responses to Demand Shocks - Crawford and Meng' Targets

	Using vendors and month specific sample average orders accepted prior to the current day as target	Using vendors, month and day-of-week specific sample average orders accepted prior to the current day as target
	(1)	(2)
Above Target	0.1810***	0.1348***
	(0.016)	(0.016)
Cumulative Orders	-0.0223***	-0.0184***
	(0.001)	(0.001)
Rainy Hour	-0.1097**	-0.1338***
	(0.047)	(0.049)
Above Target * Rainy Hour	0.0187	0.0488
	(0.047)	(0.049)
Controls	Yes	Yes
Log likelihood	-692,523.17	-86,125.657
Pseudo R^2	0.3857	0.3849
Observation	3,004,038	373,308

Notes: *** Significant at the 1% level, ** at the 5% level, * at the 10%. Standard error clustered by vendors and day are assumed. Columns (2) and (4) include dummies for vendors, day of the week, vendors and day of the week, month, year and hour of the day. The variable *Rainy Hour* takes value equal to 1 if in the corresponding hour it has rained.

Table 10: Marginal Effects on the Probability of Stopping - Target using Rainfall

	Using vendors and month specific sample average precipitation prior to the current day as target (1)	Using vendors, month and day-of-week specific sample average precipitation prior to the current day as target (2)
Cumulative Rainy Hours > Target Rainy Hours	0.0007	0.0012
	(0.003)	(0.003)
Cumulative Rainy Hours	-0.0152***	-0.016***
	(0.005)	(0.005)
Controls	Yes	Yes
Log likelihood	-694,983.98	-86,337.373
Pseudo R^2	0.3835	0.3834
Observation	3,005,038	373,308

Notes: *** Significant at the 1% level, ** at the 5% level, * at the 10%. Standard error clustered by vendors and day are assumed. Columns include dummies for vendors, day of the week, vendors and day of the week, month, year and hour of the day.

Table 11: Marginal Effects on the Probability of Stopping - Target using Weighted Rainfall

	Using vendors and month specific sample average precipitation prior to the current day as target (1)	Using vendors, month and day-of-week specific sample average precipitation prior to the current day as target (2)
Cumulative Weighted Rainy Hours > Target Weighted Rainy Hours	-0.002	-0.006
	(0.002)	(0.002)
Cumulative Weighted Rainy Hours	-0.143***	-0.190***
	(0.065)	(0.065)
Controls	Yes	Yes
Log likelihood	-694,990.47	-86,337.829
Pseudo R^2	0.3835	0.3834
Observation	3,005,038	373,308

Notes: *** Significant at the 1% level, ** at the 5% level, * at the 10%. Standard error clustered by vendors and day are assumed. Columns include dummies for vendors, day of the week, vendors and day of the week, month, year and hour of the day.

Appendix A4: First Stage Results

Table 12: First Stage Results - Dependent Variable: Crawford and Meng's Target

	Using vendors and month specific sample average orders accepted prior to the current day as target		Using vendors, month and day-of-week specific sample average orders accepted prior to the current day as target	
	Above Target	Above Target	Above Target	Above Target
Cumulative Rainy Hours > Target Rainy Hours	0.706**		0.172***	
	(0.033)		(0.040)	
Cumulative Rainy Hours	0.016**		-0.000	
	(0.006)		(0.007)	
Cum. Weight. Rainy Hours > Target Weight. Rainy Hours		0.094***		0.145***
		(0.033)		(0.034)
Cumulative Weighted Rainy Hours		0.160*		0.068
		(0.085)		(0.093)
Controls	Yes	Yes	Yes	Yes
Log likelihood	-53,572.763	-53,597.279	-48,886.782	-48,871.595
Observations	143,440	143,440	120,190	120,190
R-squared	0.3397	0.3394	0.2923	0.2925

Notes: *** Significant at the 1% level, ** at the 5% level, * at the 10%. Standard error clustered by vendors and day are assumed. Columns include dummies for vendors, day of the week, vendors and day of the week, month, year and hour of the day.

Table 13: First Stage Results - Dependent Variable: Cumulative orders

	Using vendors and month specific sample average orders accepted prior to the current day as target		Using vendors, month and day-of-week specific sample average orders accepted prior to the current day as target	
	Cum. Orders	Cum. Orders	Cum. Orders	Cum. Orders
Cumulative Rainy Hours > Target Rainy Hours	-0.199		-0.161	
	(0.161)		(0.177)	
Cumulative Rainy Hours	0.172***		0.165***	
	(0.041)		(0.043)	
Cum. Weight. Rainy Hours > Target Weight. Rainy Hours		0.129		0.201
		(0.167)		(0.151)
Cumulative Weighted Rainy Hours		1.584***		1.444***
		(0.551)		(0.525)
Controls	Yes	Yes	Yes	Yes
Observations	143,443	143,443	120,210	120,210
R-squared	0.7055	0.7051	0.7179	0.7174

Notes: *** Significant at the 1% level, ** at the 5% level, * at the 10%. Standard error clustered by vendors and day are assumed. Columns include dummies for vendors, day of the week, vendors and day of the week, month, year, and hour of the day.

Out of sight, out of mind: voting behaviour and random seating arrangements.

Alessandro Saia*

Abstract

For nearly one hundred years seats in the Chamber of the Icelandic Parliament have been allocated randomly. Using the entire universe of votes of the Icelandic Parliament from December 1995 until March 2007, I investigate how the voting behavior of Members of Parliament is influenced by the Members seating nearby. By exploiting the random seating arrangements in the Assembly Hall, I show that being seated next to Members of a different party increases the probability of not being aligned with one's own party. Using the exact spatial orientation of the peers, I provide evidence that supports the hypothesis that interaction is the main channel that explain our results.

*Contacts: alessandro.saia@unibo.it.

1 Introduction

Economists and political scientists alike, have devoted significant effort to investigate the determinants of Members of Parliament' voting behavior (Kau *et al.*, 1982; Levitt, 1996; Snyder and Groseclose, 2000; Hix, 2002; Clinton *et al.*, 2004; and Lee *et al.*, 2004). Every day, Members of Parliament (MP) face several decisions to make, some of which have a major social and economic impact. Even if their knowledge can be limited, politicians are required to vote cast their vote on a variety of subject areas. The crucial question, then, is: how do members of Parliament vote?

Several studies have provided evidence of the fact that colleagues' input is among the most important determinants of MP's voting decision. Using roll calls taken from 1807 to 1829 of the nascent in American Congress, Young (1966) suggested that living in the same boardinghouse was a dominant determinant of MPs' voting behavior. More recently, Masket (2008), using data of the California Assembly from 1941 to 1975, provided evidence of the fact that legislators sitting next to each other influence each others' voting behavior. Cohen and Malloy (2010) obtained similar results, providing evidence of the strong impact of seat location networks on the voting behavior of US Senators.

However, these studies document a series of suggestive empirical regularities, but they do not provide causal evidence over the role of colleagues' influence in MP' voting decision. Identification is challenging for two main reasons. First, individuals constantly affect each other, making the estimation extremely difficult to identify (Manski, 1993). Second, is extremely difficult to conceive a source of exogenous change in the level of interaction among Members of Parliament.

In this paper I investigate the extent to which interactions between politicians play a role in influencing MPs' voting behavior in the context of the Icelandic Parliament. I overcome identification issues exploiting a unique feature of the Icelandic Chamber: for nearly one hundred years, seats in the Icelandic Parliament have been allocated randomly. Unlike other Parliaments, where seats are assigned on the basis of political membership, the seat distribution in the Icelandic Chamber is decided by a lottery system. Using the entire universe of voting procedures held between December 1995 and March 2007, I will exploit the random seating arrangement to provide evidence for the important role of interaction spillovers in MPs' voting behavior. The results suggest that, a Members of Parliament seated next to Members of a different party are less likely to vote in alignment with their own party. I argue that the random seating arrangements, by forcing politicians to interact with Members of different parties, might facilitate a more informal exchange of ideas. To provide support for the interaction explanation, I exploit the exact spatial orientation of MP and her peers in the Chamber. The results show that MP's voting behavior is affected by Members of other parties located in the seats in the same row (right and left) but not in different rows (front and back).

The remainder of this paper is structured as follows. The next section describes the institutional background of the Icelandic Parliament. Section 3 describes the data on votes and on seating plans used in the paper. Section 4 presents the results from the empirical analysis. The mechanism through which MPs' influence operates is explored in Section 5. Section 6 concludes.

2 Institutional Background

The Althingi is the national Parliament of Iceland. It was fully established in 930 and is one of the oldest extant parliamentary institutions in the world. It exercises legislative power, jointly with the President of the Republic, and the legislative activities are held in the Parliament House, at Austurvöllur in Reykjavík. The 63 Members of the Althingi are elected by direct universal suffrage through a proportional list system. Deputies are appointed for a 4-year Legislature, which is divided in 4 sessions. Each session normally starts on October 1 and lasts for one year.

Unlike other Parliaments of the world, where seats are assigned on the basis of political membership, the seat distribution in the Icelandic Parliament is decided by a lottery system[1].

> *"At the commencement of each session Members draw lots for their seats, with the exception of the ministers who sit in ministerial chairs. When the drawing of seats begins the Speaker has on his desk a list of members, alphabetically arranged by first name. Each member is also given a number, with the member at the top of the list given the number one. Officials sit on both sides of the Speaker and they have in front of them a box of balls. Each ball is marked with a number which corresponds to a particular seat in the chamber. The Speaker starts the procedure by drawing one ball from the box on his left side. The number of that ball decides which member will be the first to approach the Speakers Chair and pick a ball from the box on the Speakers right. The Speaker then calls other members to the Chair in an alphabetical order, starting with the member whose name was first drawn. When a member has drawn a numbered ball the Speakers announces the number of the seat he has been allocated. Members take their seat accordingly. The officials record the*

[1] The random seat assignment process has been implemented in Iceland for nearly 100 years. Similar arrangements were implemented in the Philippines (between 1907 and 1988) and in the House of Representatives of the United States, from 1845 until 1913.

drawing. When a member has been assigned a seat he will retain that seat for the remainder of the legislative session. Drawing of seats has never applied to ministers, who sit in ministerial chairs, facing the assembly."

Magnusson, T. (2014)

3 Data

3.1 Voting record

Votes in Althingi had been electronically recorded since October 1991. Members of the Parliament cast their vote using the 'yes', 'no' or 'abstain' buttons on their desks. I collect all of the votes in the Icelandic Parliament between December 1995 and March 2007. During that period, a total of 23,547 voting procedures were held and 1,236,326 individual votes were recorded. Table 1 presents the distribution of votes observed (*Yes*, *No*, *Absteined*, *Absence*[2] and *Procl. Absence*[3]).

[Table 1]

High absence levels are due to the fact that Icelandic Parliament may not pass any resolution unless half the Members are present and participate in the vote, then the quorum requirement can induce strategic abstention (i.e. MP tries to spoil the outcome rather than going to the voting procedure to lose the vote).

The term *party line* $vote_{v,p}$ refers to the vote[4] that receives a substantial majority of votes cast from political party p in each voting session v. I use the term $Non\text{-}compliance_{i,v,t}$ when the vote of MP i, from political party p is different from the corresponding *party-line* $vote_{v,p}$.

Table 2 reports the distribution of votes observed in voting procedure where it was possible to identify a party line vote[5]. Table 3 and Table 4 contain the distribution of party line votes and the distribution of non-complied votes (i.e. when MP's vote is different from the *party-line vote*), respectively[6].

[2] *Absence* is recorded when the MP did not press the button during the vote (i.e. she might not be in the chamber or she was in the chamber and missed the chance to press a button).

[3] *Procl. Absence* is recorded when the MP was not able to attend the meeting (e.g. she might be overseas on parliamentary business, she might be ill or her child might be sick).

[4] i.e. *Yes*, *No*, *Absteined*, *Absence*.

[5] In the baseline sample I only consider political parties with at least 8 members. The threshold used to define the *party line* $vote_{v,p}$ is 70%. Results obtained with different cohesion thresholds or group size are essentially the same.

[6] The large majority of non-complied votes are composed by *absences*. In this sense absences can be considered as the product of a strategic calculus and to be absent from the floor at the time of the vote is a method of opposing the party position. Previous papers have provided evidence that MPs who oppose the position taken by their party generally will leave the floor at the time of the vote (Jones and Hwang,

[Table 2]

[Table 3]

[Table 4]

3.2 Seating plans

Seating plans for each voting session were obtained from the Althingi's office, based on electronic records. I use the term *peer group of MP i* to indicate the MPs located on her left, right, front and back[7]. For each MP, I compute the proportion of peers from a different party located next to her. I use the term *Peer Group Heterogeneity* to indicate the ratio between the number of peers from a different party and the total number of peers seated nearby.

To give just one example, Figure 2 shows the *peer group* of MP Hannesson during the Legislative Assembly 2014-2015. MP Hannesson has 4 peers: one seated in front (MP Neilsson), one seated behind (MP Gunnarsdòttir), one seated on the left (MP Arnadottir) and one seated on the right (MP Poroarson). Two peers out of four are members of a political party that is different from Hannesson's party. Therefore, during the legislative Assembly 2014-2015, the *Peer Group Heterogeneity* index for MP. Hannesson takes value equal to 0.5.

[Figure 1]

[Figure 2]

In the empirical analysis I focus on the voting behavior of the 57 MPs who sit in the Assembly Hall. I exclude the Prime Minister and the Ministers, who sit in ministerial chairs and for whom drawing of seats has never applied.

2005).

[7]Not all MPs have the same number of *peers*, since the number of peers depends on the seat. The number of peers varies between 2 and 4.

4 Empirical Strategy and Results

I investigate the extent to which being seated next to Members of a different party increases the probability of not being aligned with one's own party using the following model:

$$Non-compliance_{iv} = \alpha_i + \beta_1 Peer\ \ Group\ \ Heterogeneity_{iv} + \epsilon_{it} \quad (1)$$

where *Non-compliance* is a dummy that takes value equals to 1 when the vote of MP i in voting procedure v is not aligned with one's own party and *Peer Group Heterogeneity* represents the proportion of MPs of a different party located next to her[8].

[Table 5]

Table 5 presents the estimates of the reduced form model. The first columns contains only peer group heterogeneity, number of peers and MP fixed effects. The estimated effect of the variable indicating the degree of peers heterogeneity is positive and statistically significant. Columns (2), (3) and (4) include dummies for legislative session, day-of-the-week, month, year and size of parliamentary group of MP i. The size and significance of the effect remain unaffected when additional controls are included. The positive coefficient of *Peer Group Heterogeneity* suggests that being seated next to members of a different party increases the probability of not being aligned with one's own party by around 4%. Similar results are obtained with different cohesion thresholds (Table 6) or group size (Table 7).

[Table 6]

[Table 7]

These results suggests that MPs exposed to the opinion of members of other parties are less likely to vote in alignment with their own parties.

5 Potential channel

In this section I explore the channel through which peers' influence operate. There are several channels that could explain results presented in the previous section, but the most likely mechanism is ideas exchange through face-to-face interaction. The random seating

[8] As explained in the previous section, I focus on parliamentary groups p with at least 8 Members and on voting procedure v for which it is possible to identify a *party line vote* (i.e. the vote that receives the 70% of the total votes cast from party p)

arrangement force MPs to interact with members of other parties and might promote a more informal exchange of ideas.

It is reasonable to think that the effect might not be instantaneous and the effect of peers' influence might evolve through time. To explore the stability of the effect of peers' influence over time, I performed a rolling regression exercise. Figure 3 plots the coefficient of the variable *Peer Group Heterogeneity* obtained by adding one day to the sample each time: for example, the 100 data point shows the coefficient estimate based on voting data from the first day to 100 days after the beginning of the legislative session, the 200 data point shows the coefficient estimate based on data up to 200 days after the beginning of the legislative session etc. In the first part of the legislative session the effect of peer group heterogeneity is not significant statistically. At the beginning of the legislature the fact of being seated next to members of a different party doesn't affect MP's voting behavior. As time goes by the coefficient of interest starts to become positive and significant. This pattern is consistent with the hypothesis that the effect of peers' influence is generated through repeated exposure and interactions over time.

[Figure 3]

To provide further support for the interaction explanation, I exploit the spatial orientation of MP and her peers in the Chamber. I argued that the main channel through which peer group heterogeneity influences MPs' voting behavior is the face-to-face interaction. Since it is easier mechanically for a MP to interact with the members of other parties located in the seats in the same row (right and left) than those in different rows (front and back), then the role of peers seated in the same row should be different from the role of peers seated in front and in the back. To test this hypothesis I will exploit the exact spatial orientation of the peers using the following model:

$$Non-compliance_{iv} = \alpha_i + \beta_1 Left-Right\ \ Heter._{iv} + \beta_2 Back-Front\ \ Heter._{iv} + \epsilon_{it} \quad (2)$$

where the variable *Left-Right Heterogeneity* indicates the proportion of MP of a different party seated to the left and to the right and *Back-Front Heterogeneity* those seated in the front and behind.

[Table 8]

Corresponding estimates are presented in Table 8. Consistent with our hypothesis the variable *Left-Right Heterogeneity* is always positive and statistically significant while the variable *Back-Front Heterogeneity* is not. These findings provide further support for the idea that face-to-face interaction among MPs of other parties is the main mechanism that explain why politicians seated next to Members of different parties are less likely to vote in alignment with their own party.

6 Conclusion

It is generally taken for granted that seating in Parliaments are arranged on the basis of party membership and Members of Parliaments are grouped together along partisan lines. The Icelandic case, however, is very different. For nearly one hundred years seats in the Chamber of the Icelandic Parliament have been determined by lottery. Using a novel dataset on the universe of votes held from December 1995 until March 2007, I investigate how the voting behavior of Members of Parliament is influenced by the Members seating nearby. The results show that being seated next to Members of a different party increases the probability of not being aligned with one's own party by around 4%. Moreover, using the exact spatial orientation of the MPs in the Chamber, I show that MP's voting behavior is affected by Members of other parties located in the seats in the same row (right and left) but not in different rows (front and back). This result is consistent with the hypothesis that face-to-face interaction is the main mechanism responsible for our findings.

References

Clinton, J., Jackman, S., and Rivers, D. (2004). The statistical analysis of roll call data. *American Political Science Review*, 98(02):355–370.

Cohen, L. and Malloy, C. (2010). Friends in high places. *NBER Working Paper Series*, No. 16437.

Hix, S. (2002). Parliamentary behavior with two principals: preferences, parties, and voting in the european parliament. *American Journal of Political Science*, pages 688–698.

Jones, M. P. and Hwang, W. (2005). Party government in presidential democracies: Extending cartel theory beyond the us congress. *American Journal of Political Science*, 49(2):267–282.

Kau, J. B., Keenan, D., and Rubin, P. H. (1982). A general equilibrium model of congressional voting. *The Quarterly Journal of Economics*, pages 271–293.

Lee, D. S., Moretti, E., and Butler, M. J. (2004). Do voters affect or elect policies? evidence from the us house. *The Quarterly Journal of Economics*, pages 807–859.

Levitt, S. D. (1996). How do senators vote? disentangling the role of voter preferences, party affiliation, and senator ideology. *The American Economic Review*, pages 425–441.

Magnusson, T. (March 2014). Seating arrangements in the althingi. *Plenary Minutes Geneva Session, Association of Secretaries General of Parliaments.*

Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *The Review of Economic Studies*, 60(3):531–542.

Masket, S. E. (2008). Where you sit is where you stand: The impact of seating proximity on legislative cue-taking. *Quarterly Journal of Political Science*, 3:301–311.

Snyder Jr, J. M. and Groseclose, T. (2000). Estimating party influence in congressional roll-call voting. *American Journal of Political Science*, pages 193–211.

Young, J. S. (1986). *The Washington Community, 1800-1828.* Columbia University Press.

Tables and Figures

Table 1: Distribution of Recorded Votes - Full Sample

Type	Freq.	Percent
Absence	225,606	18.27
Abstained	71,803	5.81
No	54,575	4.42
Procl. Absence	87,245	7.06
Yes	795,675	64.43
Total	1,236,326	100.00

Table 2: Distribution of Recorded Votes - Group size>7 - Cohesion Rate > 70%

Type	Freq.	Percent
Absence	82,005	14.88
Abstained	30,523	5.54
No	31,405	5.70
Procl. Absence	36,243	6.58
Yes	370,804	67.30
Total	550,980	100.00

Table 3: Distribution of Party Line Votes - Group size>7 - Cohesion Rate > 70%

Type	Freq.	Percent
Absence	1,177	0.21
Abstained	36,338	6.60
No	37,967	6.89
Yes	475,498	86.30
Total	550,980	100.00

Table 4: Distribution of Non-Complied Votes - Group size>7 - Cohesion Rate > 70%

Type	Freq.	Percent
Absence	81,417	67.16
Abstained	1,487	1.63
No	426	0.41
Procl. Absence	36,243	29.90
Yes	1,087	0.90
Total	121,223	100.00

Figure 1: Seating Plan - Legislative Assembly 2014-2015 .

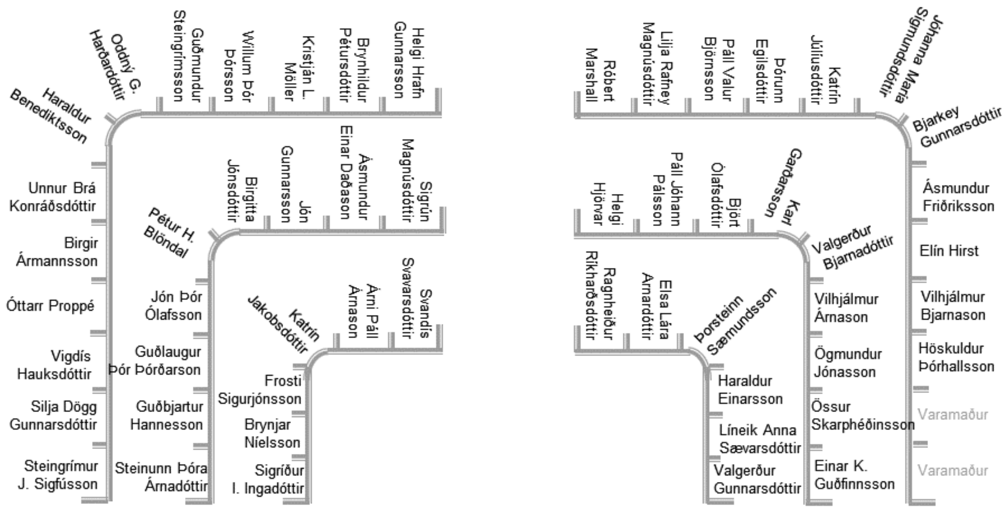


Figure 2: Detail Seating Plan - Peer Group of MP *G. Hannesson*

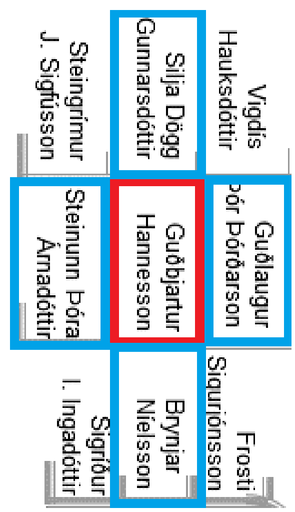


Table 5: Probability of non-compliance

	Non-compliance	Non-compliance	Non-compliance	Non-compliance
Peer Group Heterogeneity	0.0330*	0.0365**	0.0370**	0.0366**
	(0.017)	(0.017)	(0.018)	(0.018)
Number Peers	Yes	Yes	Yes	Yes
MP ID	Yes	Yes	Yes	Yes
Legislative Session	No	Yes	Yes	Yes
Group Size	No	No	Yes	Yes
Month - Year - Dow	No	No	No	Yes
Observations	514,737	514,737	514,737	514,737
R-squared	0.0701	0.0738	0.0738	0.0802

Notes: Sample: Group size>7 - Cohesion Rate > 70%. *** Significant at the 1% level, ** at the 5% level, * at the 10%. Robust standard errors clustered by day and MP are assumed.

Table 6: Probability of non-compliance - Alternative Cohesion Thresholds

	60%*	70%	80%	90%
Peer Group Heterogeneity	0.0331**	0.0366**	0.0455**	0.139***
	(0.0167)	(0.0176)	(0.0210)	(0.0304)
Observations	673,820	514,737	280,642	99,554
R-squared	0.081	0.080	0.085	0.120

Notes: Sample: Group size > 7. *** Significant at the 1% level, ** at the 5% level, * at the 10%. Robust standard errors clustered by day and MP are assumed. Regressions include number of peers, group size variable and dummies for MP, Legislature and day of the week, month, year.

Table 7: Probability of non-compliance - Alternative Cohesion Thresholds and Group Sizes

	Group Size > 6		Group Size > 8		Group Size > 9	
	(60%)	(90%)	(60%)	(90%)	(60%)	(90%)
Peer Group Heterogeneity	0.0331**	0.139***	0.0329*	0.127***	0.0335*	0.126***
	(0.0167)	(0.0304)	(0.0172)	(0.0300)	(0.0172)	(0.0300)
Observations	673,820	99,554	629,193	92,887	626,104	92,879
R-squared	0.081	0.120	0.083	0.123	0.083	0.122

Notes: *** Significant at the 1% level, ** at the 5% level, * at the 10%. Robust standard errors clustered by day and MP are assumed. Regressions include number of peers, group size variable and dummies for MP, Legislature and day of the week, month, year.

Potential Channel: Spatial orientation of the peers

Figure 3: Rolling regression - Peer Group Heterogeneity Coefficient

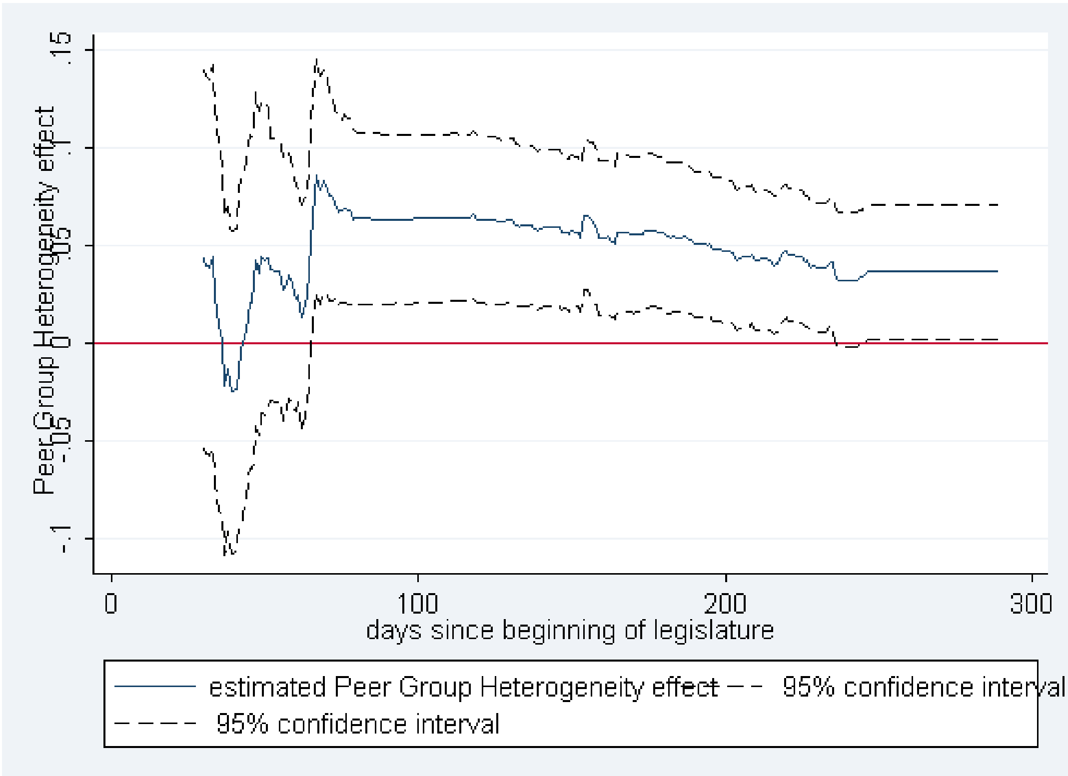


Table 8: Probability of non-compliance your own party - Spatial orientation of the peers

	Non-compliance	Non-compliance	Non-compliance	Non-compliance
Left-Right Heterogeneity	0.0305**	0.0341***	0.0342***	0.0341***
	(0.013)	(0.013)	(0.013)	(0.013)
Back-Front Heterogeneity	-0.0033	-0.0036	-0.0035	-0.0038
	(0.011)	(0.011)	(0.012)	(0.012)
Number of Peers	Yes	Yes	Yes	Yes
MP ID	Yes	Yes	Yes	Yes
Legislative Session	No	Yes	Yes	Yes
Group Size	No	No	Yes	Yes
Month - Year - Dow	No	No	No	Yes
Observations	514,737	514,737	514,737	514,737
R-squared	0.0704	0.0740	0.0740	0.0804

Notes: Sample: Group size>7 - Cohesion Threshold > 70%. *** Significant at the 1% level, ** at the 5% level, * at the 10%. Robust standard errors clustered by day and MP are assumed.

Table 9: Probability of non-compliance - Spatial orientation of the peers - Alternative Cohesion thresholds and Group Sizes

	Group Size > 6		Group Size > 8		Group Size > 9	
	(60%)	(90%)	(60%)	(90%)	(60%)	(90%)
Left-Right Heter.	0.0248**	0.01095***	0.0290**	0.0997***	0.0303**	0.0944***
	(0.012)	(0.022)	(0.012)	(0.022)	(0.012)	(0.022)
Back-Front Heter.	0.0035	0.0197	-0.0002	-0.0160	-0.0008	0.0161
	(0.011)	(0.019)	(0.011)	(0.020)	(0.011)	(0.020)
Observations	691,897	99,554	647,270	92,887	644,181	92,879
R-squared	0.081	0.120	0.084	0.123	0.084	0.123

Notes: *** Significant at the 1% level, ** at the 5% level, * at the 10%. Robust standard errors clustered by day and MP are assumed. Regressions include number of peers, group size variable and dummies for MP, Legislature and day of the week, month, year.

Choosing the open sea: The cost to the UK of staying out of the Euro

Alessandro Saia*

Abstract

In this paper we provide an estimate of the trade flows that there would have been between the UK and Europe if the UK had joined the Euro. As an alternative approach to the standard log-linear gravity equation we employ the synthetic control method. We show that the aggregate trade flows between Britain and Europe would have been 13% higher if the UK had adopted the Euro. Finally, we provide a new estimate of the Euro's trade impact for European countries. Our results suggest that the adoption of the single currency led to an increase of intra-European trade flows between 25% and 53%.

*Contacts: alessandro.saia@unibo.it. I am particularly grateful to Andrea Ichino and Tommaso Nannicini. I also received useful comments from seminar participants at Bologna and conference participants at the RIEF Doctoral Meeting and ETSG Leuven.

1 Introduction

Each time we must choose between Europe and the open sea,
we shall always choose the open sea.
Winston Churchill

From an economic point of view, the decision to become a member of the Euro was a gamble. Whereas the costs were known, the benefits were not. While the loss of the monetary channel as a tool to stabilize the economy was certain, economists have yet to decide on the benefits in terms of increased trade and, since its adoption in 1999, the Euro's effect on trade flows has been one of highly-debated question in international macroeconomics.

From an empirical point of view the gravity equation constitutes the workhorse in most applied works[1]. While the sound theoretical foundation and the strong empirical explanatory power have lead to its rapid adoption, its ability to identify the effects of currency unions on trade have displayed several problems. The fragility of the coefficient estimates, that have been characterized for their sensitiveness to the specific sample composition, and the failure to address endogeneity issues have lead to exploit alternative techniques.

When selection into treatment is not random non parametric techniques offer a valid alternative to the gravity equation[2]. By comparing the difference between trade flows of country pair in the currency union (treated unit) and similar untreated unit (counterfactual unit) it is indeed possible to estimate the causal effect of currency union on trade. Obviously, these estimates can be seriously distorted if the treated unit and its potential counterfactual are intrinsically different and reliability of the results ultimately depends on the selection of a credible counterfactual. The key is to select a group of country pairs, outside the currency union, that is able to replicate the trade flows among pairs that have adopted the same currency.

By studying the impact of Euro on trade a natural question to ask is whether trade flows between members and non-members would have been higher if non-members had joined the currency union in 1999.

In this paper we study the effect of adopting the Euro on trade flows by looking at one country that has decided not to join the currency union: the United Kingdom[3].

[1]The first attempt to investigate the Euro's impact on trade was that of Micco et al. (2003), who put the Euro effect at about 4% - 16%. Several authors have offered estimates of the effects of the common European currency. Bun and Klaassen (2007), Baldwin and Nino (2006), Flam and Nordström (2006), Frankel (2010) and Bergin and Lin (2012) provide positive estimates, while Berger and Nitsch (2008) and Silva and Tenreyro (2009) find virtually no significant effect. In Frankel (2010), the author argued, that the results may well depend on the chosen sample size: the larger the sample, the higher the resulting estimate of the currency union effect and he puts the Euro effect between 15% and 140%. This pattern appears similar to that found by Bergin and Lin (2012). Typically findings should not be interpreted as estimates of the causal effect of Euro, since they fail to address endogeneity concerns. Baldwin and Taglioni (2007) and Baldwin et al. (2008) provide a broad summary of the literature.

[2]Persson (2001) and Alesina et al. (2003) provide an excellent discussion of the problem of self-selection into currency union. Recently, several influential papers have successfully adopted nonparametric matching techniques as an alternative approach to log-linear gravity equations to investigate the causal effect of currency union (Persson (2001), Chintrakarn (2008)), free trade agreement (Egger et al. (2008), Baier and Bergstrand (2009)) and other policies aimed at promoting trade (Baghdadi et al. (2013), Chang and Lee (2011)).

[3]We know of other studies that share a similar approach. Hashem Pesaran et al. (2007) provide an

When Gordon Brown was Chancellor of the Exchequer, he claimed that "with Britain in the Euro, British trade with the Euro area could increase substantially - perhaps to the extent of 50 per cent over 30 years" (Brown (2003)). We investigate Gordon Brown's prediction by estimating how much trade would have been if the UK had joined the Euro. In doing so we will adopt a recent data driven procedure that provides a transparent way to choose counterfactual units in comparative case studies: the synthetic control method[4].

Instead of just comparing the unit of interest to a similar unit, the synthetic control method provides a systematic procedure to construct a counterfactual as convex combination of control units. In our case, we apply the synthetic matching to select a weighted combination of European country pairs that is able to reproduce the trade flows between each European country and Britain before the introduction of the common currency. Weights are selected using a transparent data-driven procedure in order to approximate the behavior of the unit of interest before the introduction of the Euro. The convex combination will represent the counterfactual version of our untreated unit and it will provide an estimate of the trade that would have been between the UK and European country if the former had adopted the Euro in 1999. If positive, the difference between the weighted unit and the actual unit would then suggest that UK's decision of not to adopt the single currency came with a cost.

The main advantage of running an analysis of the counterfactual scenario of the UK entry to the Euro is that the set of potential control units rather than arbitrary selected is *de facto* restricted to European countries that have adopted the Euro in 1999. Moreover, by limiting the scope of the paper to the UK, we focus on a member of the political and economic framework of the European Union, deeply integrated with others European countries well before 1999[5]. Apart from the transparent data drive procedure implemented to construct the synthetic counterfactual, the adopted statistical framework relies on weaker identification assumptions allowing for presence of confounding unobserved time-varying characteristics, where other techniques used in the literature (fixed effects and difference-in-differences) deal only with time-invariant unobservable factors[6].

The remainder of this paper is organized as follows. In Section 2 we describe the methodology. In Section 3 we present our estimates of the cost to the UK of staying of the Euro. In Section 4 we discuss the bias in the results of Section 3 due to general equilibrium implication and we exploit the flexibility of synthetic matching to provide evidence of the size and the sign of the

estimate of conditional probability distributions for output and inflation related with the UK entry to the Euro using a Global VAR approach. Micco et al. (2003) extend the scope of their initial paper focusing on the possible Euro's effect on the UK, during the period 1999-2001. Their estimates suggest that if the UK had joined the Euro, its trade flows would have been 7% percent larger in 2001.

[4] Abadie et al. (2010) and Imbens and Wooldridge (2009) discuss the advantages of this methods. In recent applications, the synthetic control method has proven to be a valid tool to assess the impact of policy related events (Abadie and Gardeazabal (2003), Acemoglu et al. (2013), Billmeier and Nannicini (2013) and Montalvo (2011)).

[5] The UK joined the European Economic Community in 1973 and the free movement of goods was one the cornerstone of the Single European Market.

[6] Furthermore since weights are restricted to be nonnegative and sum to one, the synthetic control method safeguards against the risk of parametric extrapolation. See Abadie et al. (2014) for a detailed discussion of the role of extrapolation in regression analysis.

bias. In Section 5 we assess the robustness of the previous results conducting an additional exercise, to obtain Euro's trade effect estimates for European countries. Section 6 concludes.

2 Methodology

Suppose that we observe X country pairs that adopted the Euro in 1999 and Y country pairs that have decided to staying out. For each $y \in Y$, we can identify the loss in trade due to the decision of not to adopt the single currency at time t using:

$$\eta_{yt} = TF_{yt}(€) - TF_{yt}(£) \qquad \forall \qquad t \geq 1999 \tag{1}$$

where $TF_{yt}(€)$ represents the value of trade flows[7] between two countries both adopting the Euro and $TF_{yt}(£)$ denotes the value of trade between the country pair not sharing the common currency. Since we are interested in the effect of UK's decision to not to adopt the Euro, $TF_{yt}(€)$ and $TF_{yt}(£)$ denote the trade flows between Britain and an European country if the UK would have joined the Euro and the actual value of trade, respectively. Obviously, we cannot compute η_{yt} since for any given year $t \geq 1999$ only $TF_{yt}(£)$ is observed.

In order to estimate η_{yt} we need to formulate a credible counterfactual of $TF_{yt}(€)$, namely a country pair (or a group of country pairs) that has adopted the Euro in 1999 and it is able to reproduce the evolution of trade if the untreated country pair would have been treated.

Suppose that for each country pair, the value of $TF_{\cdot t}$ can be given by the following factor model

$$TF_{\cdot t}(€) = \tau_{\cdot t} + \delta_t + \theta_t Z_{\cdot} + \lambda_t \mu_{\cdot} + \epsilon_{\cdot t} \tag{2}$$

$$TF_{\cdot t}(£) = \delta_t + \theta_t Z_{\cdot} + \lambda_t \mu_{\cdot} + \epsilon_{\cdot t} \tag{3}$$

where δ_t represent an unknown common determinant of trade, $Z_{\cdot}$ is a vector of observed covariates not affected by the intervention[8], θ_t denotes a vector of unknown parameters, $\mu_{\cdot}$ denotes a pair specific unobservable and λ_t is an unknown common factor that allows for the impact of

[7]Similarly to other papers (among others Rose (2000), Micco et al. (2003), Tomz et al. (2007), Disdier and Head (2008) and Baier and Bergstrand (2009)) our variable of interest is the average of two-way bilateral trade flows. Data on trade flows over the period 1980-2012 are obtained from IMF DOTS database.

[8]By assuming that the set of covariates is not affected by the intervention we have to rule out the possibility of changes in these variables in response to future adoption of the common currency (i.e. anticipation effects). To be as conservative as possible, all the elements of $Z_{\cdot}$ refer to the matching period prior to 1999. The final decision about the set of participating countries and the ultimate conversion rates were formally announced only in mid 1998 and the Euro was introduced, only into financial markets, on 1 January 1999 and all national currencies were physically three years later. Empirical evidence on the effect prior the Euro adoption is mixed. Recently Bergin and Lin (2012) studying the dynamic response of trade to Euro, have provided evidence of a positive effect on overall trade only one year before currency union adoption while Frankel (2010) shown that the effect on trade become significant much later. We conducted several robustness exercises with different matching periods, but results do not change substantively. These results are confined to the Appendix B.

unobservable pair specific determinants to vary with time.

Abadie et al. (2010) shows that using a vector of weights $W^* = (w_1^*, ..., w_X^*)$ such that

$$\{W^*\} = \underset{\{W^*\}}{\operatorname{argmin}} \sum_{t=1}^{T_0} \left(Z_{yt} - \sum_{i=1}^{X} w_i^* Z_{it} \right)^2 \tag{4}$$

where $\sum_{i=1}^{X} w_i^* = 1$, $w_i \geq 0$ for all $i = 1, ..., X$ and T_0 represents the period before to the adoption of the common currency, we can estimate

$$\widehat{\eta_{yt}} = TF_{yt}(£) - \sum_{j=1}^{X} w_i^* TF_{it}(€) \qquad \forall \qquad t \geq 1999 \tag{5}$$

Similarly to other matching techniques, the identification assumption of the synthetic control method is that as long as the synthetic unit provides a good approximation of the untreated country pair before 1999, any subsequent difference between the actual and the counterfacual unit should then represent the effect of the UK's decision of not to adopt the single currency.
If the matching window is large enough the convex combination would be able to resemble the structural parameters of our unit of interest and to successfully reproduce all the observed and unobserved determinants of trade for the specific untreated country pair. In order to obtain the optimal trade-off between computational costs and goodness of fit between the actual and the synthetic observations we compute the synthetic matching only considering pre-treatment bilateral flows as observed covariates[9].
By summing all UK-European country pairs we are able to provide an estimate of the aggregate cost to the UK of staying out of the Euro. More formally, we can estimate the cost on trade of the UK's decision of not to adopt the single currency over the period 1999-2012 using

$$\widehat{H_{UK,t}} = \frac{\sum_{1999}^{2012} \sum_{p=1}^{Y} TF_{pt}(£) - \sum_{1999}^{2012} \sum_{p=1}^{Y} \sum_{j=1}^{X} w_{pi}^* TF_{pit}(€)}{\sum_{1999}^{2012} \sum_{p=1}^{Y} \sum_{j=1}^{X} w_{pi}^* TF_{pit}(€)} \tag{6}$$

where $\sum_{p=1}^{Y} TF_{pt}(£)$ represents the sum of Anglo-European trade flows and $\sum_{p=1}^{Y} \sum_{j=1}^{X} w_{pi}^* TF_{pit}(€)$ represents the sum of the corresponding synthetic counterfactuals.
Following Abadie et al. (2010) and Acemoglu et al. (2013) we investigate the statistical significance of $\widehat{H_{UK,t}}$ through permutation inference. We construct confidence intervals by computing the cost to the UK of staying out of the Euro for 1,000 placebo groups, where each placebo group having the size equal to Y. Each counterfactual placebo is computed randomly drawing a treated country pair from the original control group X and constructing its synthetic counterfactual as a weighted average of the remaining control group units. Our estimate of the cost to the UK of staying out will be significant at 10% if it does not belong to the interval that contains the [5, 95] percentiles of the cost of staying out for placebo groups, eliminating from the 1000 placebo groups those for which the synthetic algorithm fail to provide a good match

[9] We perform several sensitivity analysis including additional covariates, obtaining similar results.

during the matching window[10].

3 An application of synthetic matching to trade flows

In this section we employ the synthetic control method to provide an estimate of the trade flows that there would have been between the UK and Europe if the UK had joined the Euro. Our initial sample contains 9 EMU members (Austria, Finland, France, Germany, Italy, the Netherlands, Spain, Portugal and Belgium-Luxembourg)[11] and the UK, for a total of 45 country pairs, 9 of which were untreated (i.e. each Anglo-European country pair), over a 33-year period, from 1980 to 2012.

Using the matching estimator we construct 9 counterfactual country pairs as a convex combination of 36 country pairs that have adopted the Euro in 1999. Weights are selected in order to best approximate the trade flows of the actual untreated units over the period 1980-1998[12]. Table 1 reports the evolution of bilateral trade for each UK-European country pairs and its synthetic control units[13]. The corresponding weights, obtained using the synthetic algorithm, are reported in Appendix A1.

[Table 1]

The comparison between the solid and dashed line during the matching window provides a measure of the goodness of fit reached by the synthetic control method. For each untreated country pair the estimated counterfactual is indeed able to provide a good match and the two series are fairly similar over the entire pre-treatment period. From 1999 the dashed line shows the evolution of trade flows if the UK had joined the Euro. In most cases the synthetic observations closely resemble untreated units well beyond the treatment date. The two lines start to diverge substantially just after the introduction of the Euro in physical form in 2002 and the synthetic counterfactuals start to outperform the actual observations.

While the ability of counterfactual units to reproduce untreated units it is common among country pairs, the treatment period displays substantial heterogeneity in terms of gap size

[10]We arbitrarily set a goodness-of-fit cutoff equal to AVG($\widehat{\tau_{\cdot t}}$) $\pm \sqrt{2}\ \sigma(\tau_{\cdot t})$, where AVG($\cdot$) represents the average effect during the pre-treatment period for the 9 untreated country pairs and $\sigma(\cdot)$ is the relative standard deviation. It exists a trade-off between number of eligible placebo counterfactuals and goodness of fit during the matching windows (the lower the cutoff, the higher is the similarity during the pre-treatment between unit of interest and placebo units). We have conducted a variety of robustness exercises using different thresholds for the cutoff rule and our results do not change substantively.

[11]In order to have a common treatment date we included only those countries that adopted the Euro in 1999. Disaggregated trade figures are not available for Luxembourg and Belgium. We have excluded Ireland due to its specific nature of the Anglo-Irish trade (see Thom and Walsh (2002)). It should be pointed out that our results are very similar whether we include Ireland. These results are confined to the Appendix B.

[12]Similarly to other papers adopting matching estimators (Egger et al. (2008)) pre-treatment period should be at least as larger as the post-treatment period. We have conducted a variety of robustness exercises using different samples over various time periods. Our results do not change substantively.

[13]In each graph the dashed line represents the weighted sum of treated country pairs. For aggregate figures, the dashed line represents the sum of each synthetic counterfactual.

among units. Table 2 reports country pair specific results over the period 1999-2012. While our estimates indicates that the adoption of Euro would have led to an increase in trade flows, final results may vary somewhat in terms of absolute magnitudes.

[Table 2]

For trade flows between Britain and Finland, France, Italy, Spain and Portugal estimated counterfactuals greatly outperform actual observations. Estimated impacts for these country pairs suggest that if the UK had joined the common European currency at the initial stage, trade between these countries would have been substantially larger. The estimated *loss*, in terms of unrealized trade, for the country pairs Britain-Finland and Britain-Portugal is of 45% and 36%, respectively. For Italy and France, the actual trade with Britain is about one-fifth lower than the synthetic counterfactual, while for Spain the loss in terms of missing trade is lower (around 10%)

For the combinations between the Netherlands, Belgium-Luxembourg, Germany and Austria synthetic units closely resemble the evolution of actual trade flows well beyond the matching window. Differences between the two series are always negative but smaller than those of first group.

By summing each untreated units and their synthetic counterfactuals we can derive estimates for the aggregate trade between Europe and the UK. In Figure 1 we compare the evolution of the actual and counterfactual aggregate trade (corresponding results are reported in Table 3).

[Figure 1]

[Table 3]

The synthetic unit closely tracks the evolution of the actual untreated unit until 2002 and the subsequent difference between the two lines indicates the positive potential effect of the adoption of the Euro.

Over the period 1999-2012 the cost, in terms of missing trade, to the UK of staying out of the Euro is around 13%. Our estimates suggest that the Euro would have started to have strongly significant effects on trade only after the physical introduction of notes and coins in 2002. During the period 1999-2001, when Euro was de facto an invisible currency only used for accounting purposes, the difference between our synthetic unit and the aggregate trade flows is small and not statistically significant. From 2002 to 2012 the effect is sizable, suggesting that the Euro would have increased the trade flows between the UK and European countries by around 15%.

We examine the robustness of our estimates investigating alternative compositions of our control group and different sets of trade predictors.

Given the relatively small number of potential control units a valid concern is the fact that our results are sensitive to the exclusion of any particular country pair. Following Abadie et al. (2010) we test the sensitivity of our results to change in our control group. We have reproduced the estimated gap between the aggregate trade between Britain and Europe and each synthetic counterfactual, obtained iteratively re-estimating the untreated units and omitting in each iteration a treated unit from the original control group. Our findings are satisfactorily robust and changes in the control group have negligible impact on our estimates.

Our results are robust to the inclusion of additional variables in the distance minimization problem. We perform the synthetic matching adding a set of covariates theoretically motivated by the multilateral resistance terms identified in Anderson and van Wincoop (2003). We follow strictly Baier and Bergstrand (2009), that have identified the sum of logs of country pair GDP, the log of bilateral distance and adjacency dummy as covariates for conducting their matching estimation. The inclusion of this additional set of covariates leaves unchanged our estimates, suggesting that our synthetic counterfactual is indeed able to replicate the structural parameters governing bilateral trade flows for untreated country pairs.

A point worth noting is that, like any other matching technique, synthetic control method relies on the so-called stable unit treatment value assumption (SUTVA). That is, we assume that the common currency only affects trade flows among Euro members and it does not influence untreated country pairs. The violation of this assumption may arise from various sources and would generate large biases in our estimates. From the previous figures it emerges that both actual and synthetic trade shown an increase in their trends after 2002. This pattern tends to suggest that the Euro might have affected also untreated country pairs, causing a violation of the no interference assumption.

Like any other nonparametric estimators, synthetic control method is unable to perform general equilibrium comparative statics. To capture general equilibrium effects a solution is to identify a system of equations that is able to capture the structural parameters of (world) trade flows. However, the number of assumptions required by this approach is what a nonparametric estimator is trying to avoid.

In the next section we will exploit the flexibility of the synthetic control method to provide evidence of the size and the sign of the bias in our estimates.

4 Euro's effect on trade between members and non members

Europe's currency union may have enhanced trade among its members or simply diverted trade away from third countries. It is clear that whether Euro had diverted trade away from Britain is our major concern. If Euro had led to significant trade diversion with the UK, our previous estimates should be interpreted as an upper bound of the true estimate of the cost to the UK of staying out of the Euro.

Despite the variety of previous results stemming from a wide set of empirical approaches no consensus has yet emerged on the existence of a negative (or positive) trade effect of Euro for third countries. To date, empirical estimates suggest that Euro had little or no effect on trade between members and non-members (see Rose (2000) and Frankel and Rose (2002) among others) and to the best of our knowledge there is no evidence of trade diversion related with the Euro.

If trade flows between the UK and Europe have been affected by the adoption of the common currency it also means that Anglo-European pairs were partially exposed to treatment.

In this section, we exploit the flexibility of the synthetic matching proposing an additional empirical exercise designed to derive the direction of the bias, otherwise difficult to evaluate with other nonparametric techniques.

We perform a second type of exercise by re-estimating a counterfactual version of trade flows between Europe and Britain using pairs of countries that have not adopted the Euro as potential controls. By considering bilateral flows among extra-European countries, where the no interference assumption is less likely to be violated, we are able to provide clear a evidence of both direction and size of the bias, since weighted combination of extra-European units should represent the counterfactual version of Anglo-European trade without the interference of the Euro. Any positive gap between the actual untreated trade and this second counterfactual should be interpreted against the fact that Euro has diverted trade have from UK and it would provide evidence in favour of the positive effect of Euro on trade between members and non members.

Apart from providing evidence for the bias in our estimates, this second exercise is meant to be interpreted as an additional inferential exercise based on the idea of the refutability test[14]. If currency union truly increases trade (and if is true that British trade would have been greater had the UK adopted the Euro), then it must also be true to say that trade between such countries would have been lower without the existence of the single currency. If the negative gap between actual British trade and the weighted average of European units truly accounts for the (potential) effect of the Euro, in this second exercise actual trade should outperform its synthetic counterfactual, computed as the weighted average of extra-European units.

Our second sample covers 11 extra-European countries (Australia, Canada, Denmark, Iceland, Japan, Norway, New Zealand, Sweden, Switzerland, the UK and United States) over the period 1980-2012[15]. Similarly to the previous section we construct the 9 Anglo-European pairs as weighted average of 55 bilateral trade flows among non-European countries.

Table 4 reports the evolution of bilateral trade for the Anglo-European pairs and their synthetic control units and country pair specific results, for the period 1999-2012, are reported in Table 5[16].

[Table 4]

[14] A well-know example of refutability test is in Freeman (1984). In the paper, the author investigated the positive effects of a trade union on workers' wages. If such a positive effect actually exists, then wages should rise when a worker moves from a non-unionized job to a unionized job, and similarly wages should fall if a worker moves from a unionized to a non-unionized job.

[15] Our sample covers extra-European countries that were OECD members over the entire matching period.

[16] In the Appendix A2 we report the weighted combinations of extra-European country pairs, obtained using the synthetic algorithm

[Table 5]

Overall, observed trade flows outperform their counterfactuals, providing strong evidence against trade diversion. Similarly to the results obtained in Section 3 we can distinguish two different groups of countries. The estimates, reported in Table 5, suggest a strong trade creation effect on trade flows between Britain and the Netherlands, Belgium-Luxembourg, Germany and Austria. The trade boosting effect of the single currency on trade between the Netherlands is of 27%, 35% for Germany, Belgium-Luxembourg is of 38% and 21% for Austria. Estimated effects for country pairs France-Britain, Italy-Britain and Spain-Britain suggest a lower but positive effect, whereas for the combination between the UK and Finland and Portugal our results provide evidence of trade diversion.

In comparing these results with those obtained in the previous section we find the exact same pattern but in reverse order. There is a negative association between the estimates for (potential) Euro trade effect and the estimates for the trade creating effect. While for the bilateral trade flows between the UK and Germany the trade-creating effects of the Euro was as strong as if the UK had joined the Euro, for Portugal the result is the opposite.

[Figure 2]

[Table 6]

Figure 2 shows the aggregate bilateral trade between the UK and European countries and its synthetic counterfactual, obtained by summing all Anglo-European country pairs. During the entire matching window the dashed line closely resembles the solid line. As in the previous section, from 2002 onwards, the two trajectories diverge substantially. Whereas in the previous section the estimated gap was positive, now the actual trade greatly outperform the synthetic observation.

Estimates reported in Table 6 are very similar to the ones displayed in Table 3 but with opposite signs. The effect starts to became large and statistically significant after the introduction of the Euro in physical form. The positive gap between the actual Anglo-European trade and its synthetic counterpart represents our estimate of the trade creating effect of Euro on bilateral trade. Aggregate estimates show that the Euro has increased the trade flows between the UK and European countries by around 21% over the period 1999-2012.

The positive difference between the two series suggests that Euro had great positive effect on trade between its members and the UK. The strong evidence of trade creation also suggests that the estimates of the (potential) Euro effect for UK, presented in the Section 3, are conservative, implying that the cost to the UK of staying out of the Euro was partially reduced by the positive stimulus to Anglo-European trade associated with the introduction of the common currency.

5 Euro's trade effect on European countries

In this section, we conduct an additional empirical exercise by providing an new estimate the Euro's trade effect for 9 European countries (Austria, Finland, France, Germany, Italy, the Netherlands, Spain, Portugal and Belgium-Luxembourg), for a total of 36 country pairs. For each treated units we construct two different synthetic counterfactuals using two different control groups. The first control group is made up of 99 partially treated country pairs, that correspond to the trade flows among 9 European countries and the 11 extra-European countries used in the previous section. The second group, instead, contains the 55 extra-European country pairs used in Section 4[17].

[Figure 3]

[Table 7]

[Table 8]

We plot in Figure 3 the aggregate trade among European countries and the two synthetic counterfactuals.[18] The dotted line represents the evolution of the weighted combination of bilateral trade flows between European and extra-European countries while the dashed line is obtained using as potential controls extra-European combination only. As in our previous estimates, the gap between actual and synthetic trade starts to accumulate well beyond 1999 and the series starts to diverge when the Euro physically replaced the national currencies. From 2002 both synthetic counterfactuals start to accumulate an increasing gap relative to the treated unit: while actual trade is characterized by a superior growth path, the two synthetic units experienced a lower upward trend, that leads to an increasing difference between the series.

The two synthetic units are, however, characterized by different growth paths. The counterfactual observation obtained as weighted average of partially treated country pairs experienced an increase in trade relative to the other counterfactual unit. This pattern is perfectly consistent with the positive external effect of Euro obtained for the British case and the corresponding gap between the dotted and the continuous line in Figure 3, should be interpreted as a lower bound estimate of the true Euro effect. In this case, the estimated effect over the period 1999-2012 is approximately 25%, fairly similar to the results obtained for the UK in Section 3.

[17]Extra-European countries are Australia, Canada, Denmark, Iceland, Japan, Norway, New Zealand, Sweden, Switzerland, the UK and United States.

[18]Using as potential counterfactuals the trade flows among European countries and extra-European countries the synthetic counterfactual is not able to provide a good match during the matching window. Using the second control group the synthetic matching is not able to find a suitable counterfactual for the country pair Germany - France. This is due to the fact that the value of trade flows between Germany and France is far from the convex hull of the trade flows of the potential control units. It is worth to noting that results reported in Table 7 and 8 are very similar when the country pair Germany-France is excluded from the 36 initial treated country-pairs.

Unlike the first counterfactual, during the period before 1999 the weighted combination of extra-European country pairs is characterized by a more stable pattern, resembling the pre-treatment trajectory well beyond the treatment date. As long as the no interference assumption olds, the difference between the dashed line and the actual unit should be interpreted as the true effect of the single currency on trade among its members. In this case the point estimate of the effect for the entire post-treatment period is actually positive and large: 53%.
Taken as a whole, our results reinforce those of the previous sections and suggest that the European single currency has greatly increased not only the bilateral trade among members, but also trade flows between members and non-members.

6 Conclusion

In reducing transaction costs, currency union may enhance trade flows among members. Despite the huge amount of work, no definitive conclusion has been reached so far. In this paper we use a nonparametric matching technique to provide an estimate of the trade flows that there would have been between the UK and Europe if the UK had joined the Euro. Using the synthetic control method, we construct a counterfactual unit for bilateral trade flows between the UK and each European country, obtained as a convex combination of bilateral flows among European countries. Our results indicate that aggregate trade flows between Britain and Europe would have been 13% higher if the UK had adopted the Euro. We performed several robustness exercises, by removing units from the control group, including additional predictors and considering different matching windows, that support our initial estimates.
Since we are not able to perform general equilibrium comparative statics, we exploit the flexibility of the adopted statistical methodology to investigate the direction of the bias. We provide evidence that the introduction of the Euro had a positive impact on Anglo-European trade, implying that the actual cost of staying out of the Euro has been mitigated.
Finally, we provide a bounded estimate of the impact of Euro on trade for European countries. We find that the adoption of the single currency led to an increase of trade flows among its members that ranges between 28% and 53%. While we do also find relevant discrepancies among country specific effects, taken as a whole our results suggest that Euro has greatly increased not only bilateral trade among members, but also trade flows between members and non-members.

References

Abadie, A., Diamond, A., and Hainmueller, J. (2010). Synthetic control methods for comparative case studies: Estimating the effect of california's tobacco control program. *Journal of the American Statistical Association*, 105(490):493–505.

Abadie, A., Diamond, A., and Hainmueller, J. (2014). Comparative politics and the synthetic control method. *American Journal of Political Science*.

Abadie, A. and Gardeazabal, J. (2003). The economic costs of conflict: A case study of the basque country. *The American Economic Review*, 93(1):113–132.

Acemoglu, D., Johnson, S., Kermani, A., Kwak, J., and Mitton, T. (2013). The value of connections in turbulent times: Evidence from the united states. *NBER Working Paper*, (w19701).

Alesina, A., Barro, R. J., and Tenreyro, S. (2003). Optimal currency areas. In *NBER Macroeconomics Annual 2002, Volume 17*, pages 301–356. MIT Press.

Anderson, J. E. and van Wincoop, E. (2003). Gravity with gravitas: A solution to the border puzzle. *The American Economic Review*, 93(1):170–192.

Baghdadi, L., Martinez-Zarzoso, I., and Zitouna, H. (2013). Are rta agreements with environmental provisions reducing emissions? *Journal of International Economics*, 90(2):378–390.

Baier, S. L. and Bergstrand, J. H. (2009). Estimating the effects of free trade agreements on international trade flows using matching econometrics. *Journal of International Economics*, 77(1):63–76.

Baldwin, R. and Taglioni, D. (2007). Trade effects of the euro: A comparison of estimators. *Journal of Economic Integration*, 22(4):780–818.

Baldwin, R. E., Di Nino, V., Fontagné, L., De Santis, R., and Taglioni, D. (2008). Study on the impact of the euro on trade and foreign direct investment. *European Economic and Monetary Union Working Paper*, (321).

Baldwin, R. E. and Nino, V. D. (2006). Euros and zeros: The common currency effect on trade in new goods. *NBER Working Paper*, (w12673).

Berger, H. and Nitsch, V. (2008). Zooming out: The trade effect of the euro in historical perspective. *Journal of International Money and Finance*, 27(8):1244–1260.

Bergin, P. R. and Lin, C.-Y. (2012). The dynamic effects of a currency union on trade. *Journal of International Economics*, 87(2):191–204.

Billmeier, A. and Nannicini, T. (2013). Assessing economic liberalization episodes: A synthetic control approach. *Review of Economics and Statistics*, 95(3):983–1001.

Brown, G. (2003). 9 june 2003 statement by the chancellor of the exchequer on uk membership of the single currency. *UK House of Commons, London.*

Bun, M. and Klaassen, F. (2007). The euro effect on trade is not as large as commonly thought. *Oxford Bulletin of Economics and Statistics*, 69(4):473–496.

Chang, P.-L. and Lee, M.-J. (2011). The wto trade effect. *Journal of International Economics*, 85(1):53–71.

Chintrakarn, P. (2008). Estimating the euro effects on trade with propensity score matching. *Review of international economics*, 16(1):186–198.

Disdier, A.-C. and Head, K. (2008). The puzzling persistence of the distance effect on bilateral trade. *The Review of Economics and Statistics*, 90(1):37–48.

Egger, H., Egger, P., and Greenaway, D. (2008). The trade structure effects of endogenous regional trade agreements. *Journal of International Economics*, 74(2):278–298.

Flam, H. and Nordström, H. (2006). *Euro effects on the intensive and extensive margins of trade*. CESifo.

Flam, H. and Nordström, H. (2007). Explaining large euro effects on trade: the extensive margin and vertical specialization. *Manuscript, Institute for International Economic Studies, Stockholm University*.

Frankel, J. (2010). The estimated trade effects of the euro: Why are they below those from historical monetary unions among smaller countries? In *Europe and the Euro*, pages 169–212. University of Chicago Press.

Frankel, J. and Rose, A. (2002). An estimate of the effect of common currencies on trade and income. *The Quarterly Journal of Economics*, 117(2):437–466.

Freeman, R. (1984). Longitudinal analysis of the effect of trade unions. *Journal of Labor Economics*, 2(1).

Hashem Pesaran, M., Vanessa Smith, L., and Smith, R. P. (2007). What if the uk or sweden had joined the euro in 1999? an empirical evaluation using a global var. *International Journal of Finance & Economics*, 12(1):55–87.

Imbens, G. W. and Wooldridge, J. M. (2009). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature*, 47(1):5–86.

Micco, A., Stein, E., and Ordoñez, G. (2003). The currency union effect on trade: early evidence from emu. *Economic policy*, 18(37):315–356.

Montalvo, J. G. (2011). Voting after the bombings: A natural experiment on the effect of terrorist attacks on democratic elections. *Review of Economics and Statistics*, 93(4):1146–1154.

Persson, T. (2001). Currency unions and trade: how large is the treatment effect? *Economic Policy*, 16(33):433–462.

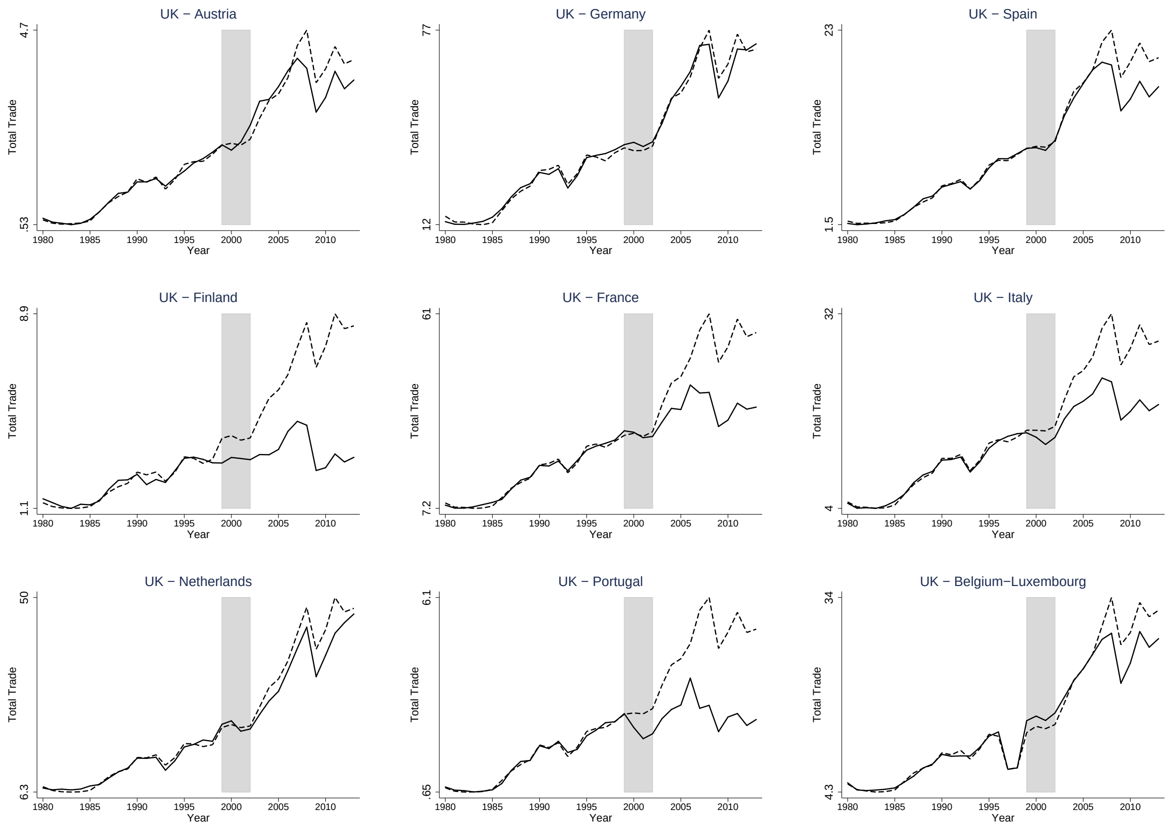
Rose, A. (2000). One money, one market: the effect of common currencies on trade. *Economic policy*, 15(30):7–46.

Silva, J. S. and Tenreyro, S. (2009). Currency unions in prospect and retrospect. *Annual Review of Economics*, 2(1):51–74.

Thom, R. and Walsh, B. (2002). The effect of a currency union on trade: Lessons from the irish experience. *European Economic Review*, 46(6):1111–1123.

Tomz, M., Goldstein, J. L., and Rivers, D. (2007). Do we really know that the wto increases trade? comment. *American Economic Review*, 97(5):2005–2018.

Table 1: Trade between the UK and European countries vs synthetic counterfactuals.



Notes: Outcome: bilateral trade (in billion of $). Solid line: actual trade between the UK and the corresponding European country. Dashed line: synthetic aggregate trade. Synthetic counterfactuals are constructed as weighted average of European country pairs using as matching window the period 1980-1998. The grey column represents the period between 1999 and 2002, a transitional period where the Euro was de facto a parallel currency.

Table 2: Trade between the UK and European countries vs synthetic counterfactuals.

	AUT	GER	ESP	FIN	FRA	ITA	NLD	PRT	XBL
% Diff. Actual vs Synth.	-5.2	-1.4	-10.1	-45.1	-24.3	-23.3	- 8.3	-36.2	- 6.0

Notes: The table reports the percentage differences between actual and synthetic aggregate trade over the period 1999-2012 for each Anglo-European country pair. Synthetic counterfactuals are constructed as weighted average of European country pairs using as matching window the period 1980-1998.

Figure 1: Aggregate trade between the UK and Europe vs its synthetic counterfactual.

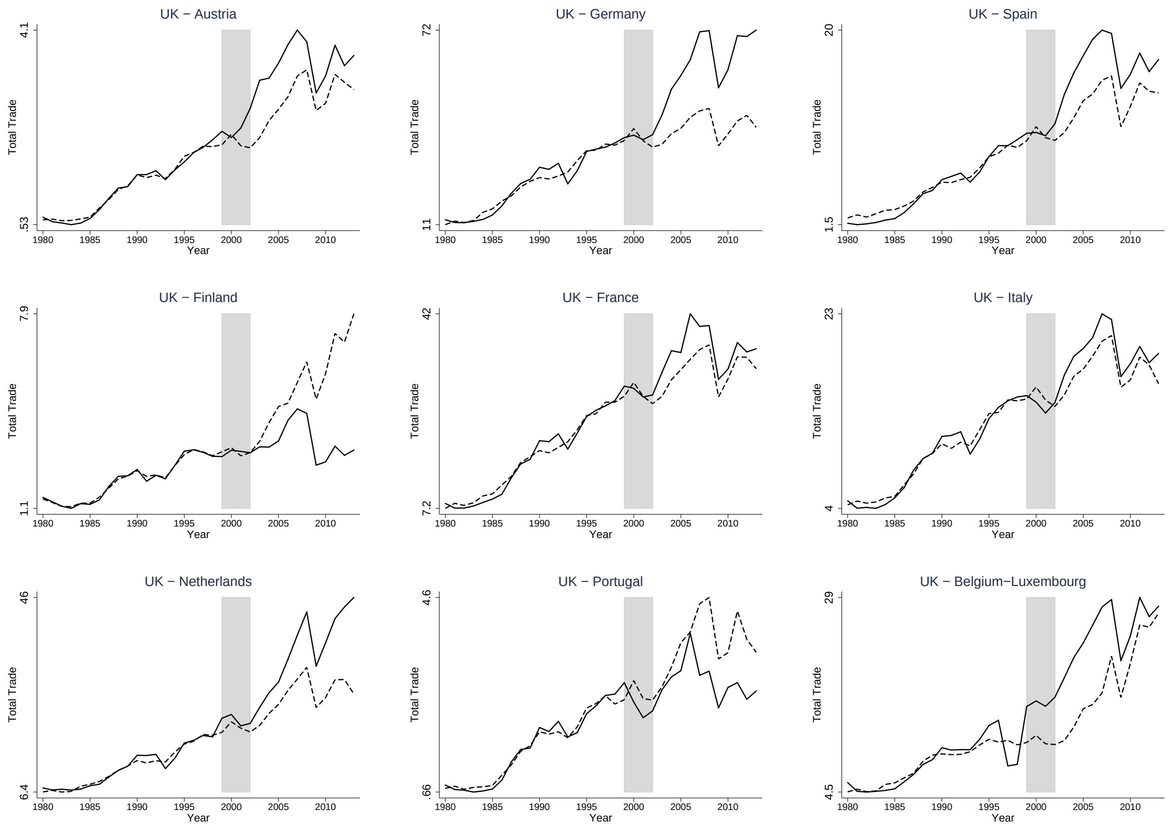
Notes: Outcome: bilateral trade (in billion of $). Solid line: actual aggregate trade between the UK and Europe (computed as the sum of trade flows between each Anglo-European country pairs). Dashed line: synthetic aggregate trade (computed as the sum of synthetic trade flows). Synthetic counterfactuals are constructed as weighted average of European country pairs using as matching window the period 1980-1998. The grey column represents the period between 1999 and 2002, a transitional period where the Euro was de facto a parallel currency.

Table 3: Aggregate trade between the UK and Europe vs its synthetic counterfactual.

	1980-1998	1999-2012		
		1999-2012	1999-2002	2002-2012
% Difference Actual vs Synth.	0.31	-12.96**	1.06	-15.17**

Notes: The table reports the percentage difference between actual and synthetic aggregate trade over the corresponding periods. Synthetic counterfactuals are constructed as weighted average of European country pairs using as matching window the period 1980-1998. Results are derived from trade flows represented in Figure 1. Confidence intervals were obtained using the method described in Section 2 (*** Significant at the 1% level, ** at the 5% level, * at the 10%).

Table 4: Trade between the UK and European countries vs synthetic counterfactuals.



Notes: Outcome: bilateral trade (in billion of $). Solid line: actual trade between the UK and the corresponding European country. Dashed line: synthetic aggregate trade. Synthetic counterfactuals are constructed as weighted average of extra-European country pairs using as matching window the period 1980-1998. The grey column represents the period between 1999 and 2002, a transitional period where the Euro was de facto a parallel currency.

Table 5: Trade between the UK and European countries vs synthetic counterfactuals.

	AUT	GER	ESP	FIN	FRA	ITA	NLD	PRT	XBL
% Diff. Actual vs Synth.	21.5	35.0	22.8	-28.6	8.9	5.7	27.2	-18.4	37.9

Notes: The table reports the percentage differences between actual and synthetic aggregate trade over the period 1999-2012 for each Anglo-European country pair. Synthetic counterfactuals are constructed as weighted average of extra-European country pairs using as matching window the period 1980-1998.

Figure 2: Aggregate trade between the UK and Europe vs its synthetic counterfactual.

Notes: Outcome: bilateral trade (in billion of $). Solid line: actual aggregate trade between the UK and Europe (computed as the sum of trade flows between each Anglo-European country pair). Dashed line: synthetic aggregate trade (computed as the sum of synthetic trade flows). Synthetic counterfactuals are constructed as weighted average of extra-European country pairs using as matching window the period 1980-1998. The grey column represents the period between 1999 and 2002, a transitional period where the Euro was de facto a parallel currency.

Table 6: Aggregate trade between the UK and Europe vs its synthetic counterfactual.

	1980-1998	1999-2012		
		1999-2012	1999-2002	2002-2012
% Difference Actual vs Synth.	-0.31	21.20***	4.14	25.03***

Notes: The table reports the percentage difference between actual and synthetic aggregate trade over the corresponding periods. Synthetic counterfactuals are constructed as weighted average of extra-European country pairs using as matching window the period 1980-1998. Results are derived from trade flows represented in Figure 2. Confidence intervals were obtained using the method described in Section 2 (*** Significant at the 1% level, ** at the 5% level, * at the 10%).

Figure 3: Aggregate trade among European countries vs synthetic counterfactuals.

Notes: Outcome: bilateral trade (in billion of $). Solid line: actual aggregate among European countries (computed as the sum of trade flows among European country pairs). Dotted line: aggregate synthetic trade computed as the sum of synthetic trade flows constructed as weighted combination European-extra-European country pairs. Dashed line: aggregate synthetic trade computed as the sum of synthetic trade flows constructed as weighted combination extra-European country pairs. Synthetic counterfactuals are constructed using as matching window the period 1980-1998. The grey column represents the period between 1999 and 2002, a transitional period where the Euro was de facto a parallel currency.

Table 7: Aggregate trade among European countries vs synthetic counterfactual.

	1980-1998	1999-2012		
		1999-2012	1999-2002	2002-2012
% Difference Actual vs Synth.	-.440***	53.02***	4.464	64.26***

Notes: The table reports the percentage difference between actual and synthetic aggregate trade over the corresponding periods. Synthetic counterfactuals are constructed using as matching window the period 1980-1998. Results are derived from trade flows represented in Figure 3. Confidence intervals were obtained using the method described in Section 2 (*** Significant at the 1% level, ** at the 5% level, * at the 10%).

Table 8: Aggregate trade among European countries vs synthetic counterfactual.

	1980-1998	1999-2012		
		1999-2012	1999-2002	2002-2012
% Difference Actual vs Synth.	8.295***	24.77***	6.001***	28.10***

Notes: The table reports the percentage difference between actual and synthetic aggregate trade over the corresponding periods. Synthetic counterfactuals are constructed using as matching window the period 1980-1998. Results are derived from trade flows represented in Figure 3. Confidence intervals were obtained using the method described in Section 2 (*** Significant at the 1% level, ** at the 5% level, * at the 10%).

Appendix A

A1. List of counterfactuals and corresponding weights selected by the SCM's algorithm to build the synthetic counterfactual (Donor pool: European country pairs)

<u>Austria - United Kingdom</u>: Austria - Spain (0.324), Austria - Portugal (0.212), Finland - France (0.22), Finland - Italy (0.156), France - Netherlands (0.036), Italy - Netherlands (0.052).

<u>France - United Kingdom</u>: Germany - Netherlands (0.139), France - Spain (0.495), France - Italy (0.294), France - Netherlands (0.072).

<u>Italy - United Kingdom</u>: France - Italy (0.171), Italy - Netherlands (0.622), Spain - France (0.128), Spain - Italy (0.079).

<u>Germany - United Kingdom</u>: France - Italy (0.476), Germany - France (0.323), Germany - Netherlands (0.036) Spain - France (0.166).

<u>Netherlands - United Kingdom</u>: France - Netherlands (0.845), France - Spain (0.058), Germany - Netherlands (0.097).

<u>Spain - United Kingdom</u>: Finland - Italy (0.144), France - Spain (0.181), Italy - Spain (0.094), Italy - Netherlands (0.273), Spain - Netherlands (0.211), Spain - Portugal (0.097).

<u>Portugal - United Kingdom</u>: Austria - Spain (0.21), Austria - Portugal (0.055), Finland - France (0.332), Finland - Portugal (0.216), France - Netherlands (0.033), Italy - Netherlands (0.073), Italy - Spain (0.047), Netherlands - Portugal (0.034).

<u>Finland - United Kingdom</u>:Austria - Finland (0.621), Belgium/Luxembourg - Finland (0.074), Belgium/Luxembourg - Netherlands (0.007), Finland - Netherlands (0.127), France - Netherlands (0.171).

<u>Belgium/Luxembourg - United Kingdom</u>: Belgium/Luxembourg - Finland (0.428), Belgium/Luxembourg - France (0.327), Belgium/Luxembourg - Netherlands (0.057), France - Netherlands (0.186), France - Spain (0.002).

A2. List of counterfactuals and corresponding weights selected by the SCM's algorithm to build the synthetic counterfactual (Donor pool: extra-European pairs)

Austria - United Kingdom: Australia - Switzerland (0.178), Canada - Norway (0.081), Canada - Switzerland (0.044), Denmark - Norway (0.03), Denmark - Sweden (0.022), Iceland - United Kingdom (0.542), Japan - United Kingdom (0.063), Switzerland - United States (0.027), United Kingdom - United States (0.014) .

France - United Kingdom: Canada - United States (0.042), Japan - United Kingdom (0.737), Japan - United States (0.052), United Kingdom - United States (0.169) .

Italy - United Kingdom: Denmark - Sweden (0.265), Japan - United Kingdom (0.572), Japan - United States (0.15), United Kingdom - United States (0.013).

Germany - United Kingdom: Denmark - Sweden (0.265), Japan - United Kingdom (0.201), Japan - United States (0.311), United Kingdom - United States (0.223).

Netherlands - United Kingdom: Australia - Japan (0.033), Canada - Japan (0.078), Denmark - Sweden (0.382) , Norway - United Kingdom (0.213), Japan - United States (0.052), United Kingdom - United States (0.242).

Spain - United Kingdom: Canada - Norway (0.727), Canada - United States (0.033), Japan - United Kingdom (0.24).

Portugal - United Kingdom: Australia - Swizerland (0.082), Canada - Norway (0.381), Denmark - Norway (0.107), Denmark - United Kingdom (0.025), Iceland - United Kingdom (0.262), Japan - Switzerland (0.04), Japan - United Kingdom (0.089), United Kingdom - United States (0.013).

Finland - United Kingdom: Australia - Canada (0.284), Australia - New Zealand (0.194), Canada - United Kingdom (0.052), Denmark - Sweden (0.067), Japan - Sweden (0.045), Sweden - United Kingdom (0.083), Sweden - Switzerland (0.132), Switzerland - United Kingdom (0.143).

Belgium/Luxembourg - United Kingdom: Australia - Japan (0.384), Canada - Japan (0.191), Japan - United States (0.033), Japan - Switzerland (0.06), Switzerland - United Kingdom (0.332).

Appendix B

B1. Alternative matching window: 1980 - 1995.

Figure 4: Aggregate trade between the UK and Europe vs its synthetic counterfactual. Control group: European country pairs.

Notes: Outcome: bilateral trade (in billion of $). Solid line: actual aggregate trade between the UK and Europe (computed as the sum of trade flows between each Anglo-European country pair). Dashed line: synthetic aggregate trade (computed as the sum of synthetic trade flows). Synthetic counterfactuals are constructed as weighted average of European country pairs using as matching window the period 1980-1995. The grey column represents the period between 1999 and 2002, a transitional period where the Euro was de facto a parallel currency.

Table 9: Aggregate trade between the UK and Europe vs its synthetic counterfactual. Control group: European country pairs.

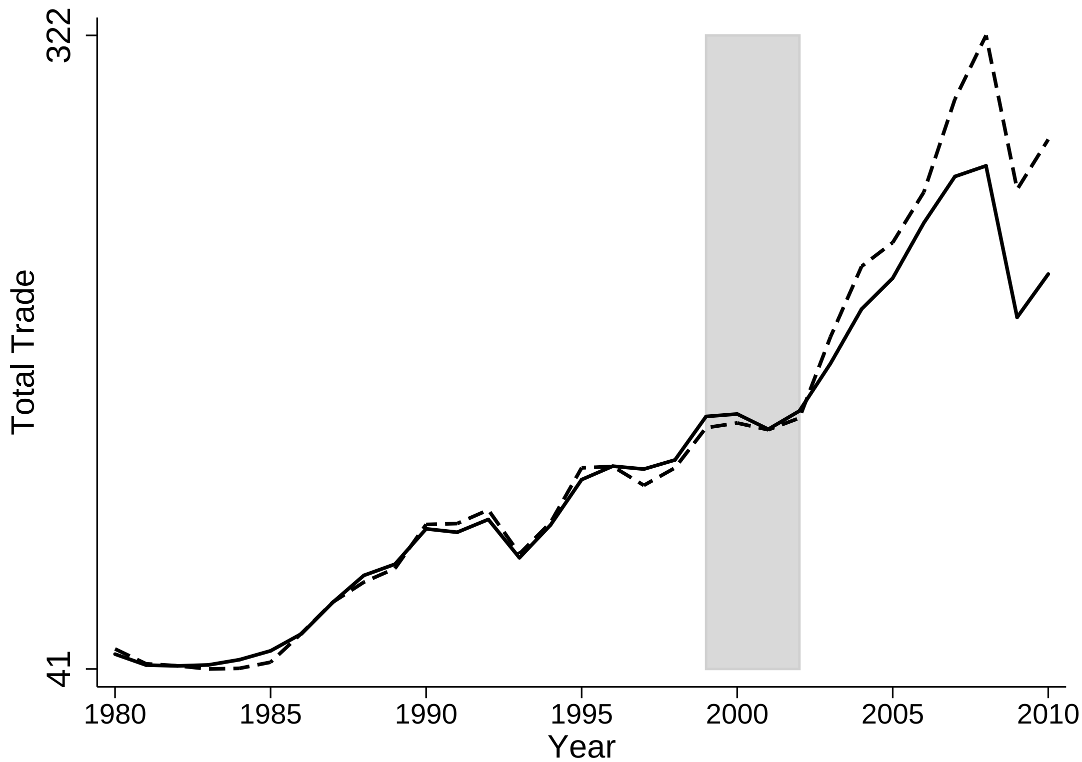
	1980-1998	1999-2012		
		1999-2012	1999-2002	2002-2012
% Difference Actual vs Synth.	1.65	-11.5**	5.46	-14.11**

Notes: The table reports the percentage difference between actual and synthetic aggregate trade over the corresponding periods. Synthetic counterfactuals are constructed as weighted average of European country pairs using as matching window the period 1980-1995. Results are derived from trade flows represented in Figure 1. Confidence intervals were obtained using the method described in Section 2 (*** Significant at the 1% level, ** at the 5% level, * at the 10%).

B2. Alternative sample composition

Potential counterfactuals: 10 European countries (Austria, Finland, France, Germany, Italy, the Netherlands, Spain, Portugal, Belgium-Luxembourg and Ireland).

Figure 5: Aggregate trade between the UK and Europe vs its synthetic counterfactual. Control group: European country pairs.



Notes: Outcome: bilateral trade (in billion of $). Solid line: actual aggregate trade between the UK and Europe (computed as the sum of trade flows between each Anglo-European country pair). Dashed line: synthetic aggregate trade (computed as the sum of synthetic trade flows). Synthetic counterfactuals are constructed as weighted average of European country pairs using as matching window the period 1980-1998. The grey column represents the period between 1999 and 2002, a transitional period where the Euro was de facto a parallel currency.

Table 10: Aggregate trade between the UK and Europe vs its synthetic counterfactual. Control group: European country pairs.

	1980-1998	1999-2012		
		1999-2012	1999-2002	2002-2012
% Difference Actual vs Synth.	.31	-9.72*	5.46	-12.1*

Notes: The table reports the percentage difference between actual and synthetic aggregate trade over the corresponding periods. Synthetic counterfactuals are constructed as weighted average of European country pairs using as matching window the period 1980-1998. Results are derived from trade flows represented in Figure 1. Confidence intervals were obtained using the method described in Section 2 (*** Significant at the 1% level, ** at the 5% level, * at the 10%).