

УДК 519.862

I. Lukianenko, Y. Gorodnichenko, S. Huzenko

## THE PROBLEMS OF THE ESTIMATION OF THE BORDER EFFECT IN UKRAINE

*This paper is the first to provide an examination of the border effect between the East and the West parts of Ukraine using the specific methodology analyzing if price differentials exhibit a border effect. We find that the estimate of border effect is generally small. The distance equivalent of the border effect does not exceed 560 km, which is small when compared to the distance equivalent of administrative borders in other countries. We also find the important determinants of the border effect. The general conclusion is that there is little economic evidence of the East-West split. There is no immediate threat of disintegration and no pressing need to transform the constitutional arrangements of powers and territories.*

### Introduction

The Presidential and parliamentary elections during the last years in Ukraine revealed a clear division into the East and the West according to voting patterns. This division has been heavily exploited by many politicians, social leaders and journalists both within the country and abroad to advocate federalization of the country. There is however very little evidence on whether the split is real or imaginary. The goal of this paper is to answer how far the West is from the East along different dimensions (economic, political, cultural and others), whether markets are segmented according to voting patterns, and whether there exists an objective risk of the division of Ukraine into two parts.

To answer these questions, we use the border effect framework. The border effect is a name for

a regularity that an administrative border between any two geographical regions is associated with reduced trade and increased price dispersion across these regions [2, 7]. It may incorporate a big range of different factors that prevent complete market integration between countries and regions within a single country. The issue of border effect is relatively new and started to draw close attention of researchers only about a decade ago. All the research works can be divided into two big subgroups of those in which quantity data and gravity-type model were used to measure border effect following McCallum and those in which border effect was found from the price data with the help of methodology introduced by Engel and Rogers [4, 5, 10, 11]. We use the methodology proposed by Engel and Rogers [4, 5]. Specifically, we examine if price differentials exhibit a border effect;

that is, price dispersion is *ceteris paribus* greater across the border than within borders of the political voting pattern. The research that we done in this paper: the investigation of the border effect in the Ukrainian context, have conducted is novel in several respects. First, all the earlier studies on the border effect tried to measure the influence of national or regional administrative borders that formally exist, whereas we apply this framework to estimate the role of the hypothetical border. Second, nobody has ever tried to find an answer to a political economy question with the help of the border effect concept. Besides, we use several modifications of the basic model, introduce some political and social variables into it, and control for the «river» effect, which has never been done yet. We look at the issue of the Eastern and Western Ukraine from an economic perspective and test whether there is also an economic East-West division of Ukraine in addition to political. Substantial differences of price volatility across East-West border might reveal this division. Finding strong economic evidence for the split would mean that it is not just a short-term temporary phenomenon and should be treated more seriously. We also estimate the role of different factors, such as relative wage volatility, gross added value per capita, political and linguistic preferences, presence of the Dnipro River in explaining the gap between the East and the West of Ukraine.

Moreover, this research has important regional policy implications. Finding a significant border effect would suggest the presence of substantial differences in tastes and preferences, levels of life, social and business networks, institutions etc. in the East and the West of Ukraine, since there are no formal trade barriers between them. It would be a signal to policy makers that they should take certain economic policy actions for bringing the East and the West together in order to avoid social tension and possible threat of separatism. It would also support a sharp need for Administrative reform in Ukraine and a deeper consideration of pros and cons of transforming.

### Theoretical background of the research

Firstly, it's necessarily shortly describe the advantages and disadvantages of the classical methodology introduced by Engel and Rogers and the modifications needed for the transitions country as well for Ukraine. In their influential paper Engel and

Rogers present a simple theoretical framework that shows the effects of distance and the border on price variation across territories, and then suggest an econometric model based on it [4, 5]. Basic assumptions behind their theoretical model are: a) all the goods have a tradable and non-tradable components, where non-tradable component might reflect, for instance, distribution and marketing costs; b) the price of tradable component of each good is determined in competitive market; c) the price of non-tradable component is set by profit-maximizing monopolist; d) Cobb-Douglas production technology with constant returns to scale.

Not all of these assumptions are very realistic, especially in the context of transition country like Ukraine. For instance, while assumption (a) seems to be equally valid both for developed and transition countries, assumption (b) is rather disputable. Even the price of a tradable component of each good does not necessarily have to be determined in competitive market. While it can be generally true for food products, this assumption is likely to be violated for nonfood products, which are usually highly differentiated, so oligopoly seems to be more appropriate for them. Besides, in the case of high capital and labor mobility arbitrage is not possible for both tradable and non-tradable components, and their prices are determined in a similar way, so assumption (c) would not generally hold either. Another problem, especially relevant for transition countries, is the state regulation of prices and state interventions in the market. In Ukraine, for instance, high level of state regulation is observed in many market of food products like sugar, bread and cereals markets. Assumption (d) is also rather restrictive: production technologies vary over industries, and there are industries with increasing returns to scale (IRS) like natural monopolies.

Despite many of its assumptions do not exactly correspond to reality, the model offered by Engel and Rogers provides some very useful insights on the factors that influence prices variation of different products across locations.

The price of good  $i$  in location  $j$  is determined according to the following formula:

$$p_j^i = \beta_j^i \alpha_j^i (w_j^i)^{\gamma_i} (q_j^i)^{1-\gamma_i}, \quad (1)$$

where  $\gamma_i$  stands for the share of non-tradable component of good  $i$  and  $w_j^i$  – for its price in location  $j$ . The share of tradable component is respectively  $(1 - \gamma_i)$  and its price in location  $j$  is  $q_j^i$ . The pro-

<sup>1</sup>Namely, results of voting during Presidential elections of 2004 and percentage of people whose native language is Russian.

ductivity is measured by  $\alpha'_j$  and the markup over costs by  $\beta'_j$  inversely related to the elasticity of demand

$$\left( \varepsilon : \beta = \frac{\varepsilon}{\varepsilon - 1} \right).$$

On the basis of theoretical predictions mentioned earlier, Engel and Rogers offer the following econometric model:

$$V(P_{jkt}^i) = \beta_1^i \ln dist_{jk} + \beta_2^i Border_{jk} + \sum_{m=1}^N \gamma_m^i D_m + u_{jk}, \quad (2)$$

where  $V(P_{jkt}^i)$  stands for price volatility measured as a standard deviation across time-series of  $P_{jkt}^i$  and

$$P_{jkt}^i = \log \left( \frac{P_{j,t}^i}{P_{k,t}^i} \right) - \log \left( \frac{P_{j,t-1}^i}{P_{k,t-1}^i} \right),$$

which shows percentage difference of relative prices of product  $i$  in locations  $j$  and  $k$  at time  $t$  and  $t - 1$ <sup>2</sup>;  $\ln dist_{jk}$  is the log of distance between locations  $j$  and  $k$ ,  $Border_{jk}$  is a dummy variable, which equals 1 when locations  $j$  and  $k$  are in different regions and 0 when they are in the same region,  $D_m$  is a dummy variable for each of  $N$  locations<sup>3</sup>,  $u_{jk}$  – regression error.

It is necessary to mention that in our research ‘location’ would stand for 25 major administrative units of Ukraine (24 oblasts and Autonomous Republic of Crimea). Two ‘regions’ would be differentiated according to political division: the East and the West, and 5 for historical: the West, the North-Center, the North-East, the South and the East.

Basic hypothesis implied by this model is that, controlling for distance, price discrepancy should be higher for locations separated by the border. Also, distance is supposed to have a positive impact on the price volatility ( $\beta_1^i > 0$ ). Coefficient  $\beta_2^i$  in this specification shows the difference between the mean price volatility of two jurisdictions located across the border and within one region after controlling for distance.

Regression (2) represents a simple OLS cross-section regression, which is run for each product separately. However, a pooled regression for all the products can be also considered because it has more observations and gives more precise estimates. Then it is appropriate to include dummies for all the products and all but one locations into regression. Pooled regression would give the average of the logged distance and border coefficients across all the goods  $\beta_1$  – and  $\beta_2$ :

$$V(P_{jkt}) = \beta_1 \ln dist_{jk} + \beta_2 Border_{jk} + \sum_{i=1}^K \lambda_i G_i + \sum_{m=1}^{N-1} \gamma_m D_m + u_{jk}. \quad (3)$$

Here  $G_i$  is a dummy variable for each of  $K$  products and  $D_u$  is a dummy variable for each but one of  $N$  locations. The rest of variables are the same as in specification (2). Natural log specification of distance is rather strong assumption, which implies a concave relationship between distance and relative price volatility. Another drawback of this specification is that this measure of distance is unitless. So, an alternative quadratic distance specification can be introduced, which would allow to test whether the assumption of concave relationship is realistic:

$$V(P_{jkt}^i) = \beta_1^i dist_{jk} + \beta_2^i sqrdist_{jk} + \beta_3^i Border_{jk} + \sum_{m=1}^N \gamma_m^i D_m + u_{jk}. \quad (4)$$

A convex specification of distance can be tried as well. In this case it is assumed that after some critical level additional distance does not influence at all relative price volatility [4].

An important issue is an economic significance of the border relative to distance in explaining price variation across locations. There are several ways to find distance equivalent of the according to the formula:

$$\exp \left( \frac{\beta_2}{\beta_1} \right),$$

where  $\beta_1$  and  $\beta_2$  are average coefficients of logged distance and border dummy respectively. However, this measure would be very sensitive to small changes in  $\beta_1$  and  $\beta_2$  because distance enters the regression in logs. Besides, under this specification interpretation of the distance equivalent would change if we change the units in which distance is measured. Parsley and Wei offer an alternative way to compute distance equivalent by finding how much more distant must be the countries (regions) in order to have the observed price dispersion [11]:

$$\beta_1 \ln(\overline{dist} + Z) = \beta_2 + \beta_1 \ln(\overline{dist}). \quad (5)$$

In the equation (5)  $\overline{dist}$  is an average distance between city-pairs across regions, and  $Z$  is actually a distance equivalent of the border effect. One can easily rearrange terms in equation (5) to solve it for  $Z$ :

$$Z = \overline{dist} * \left( \exp \left( \frac{\beta_2}{\beta_1} \right) - 1 \right). \quad (6)$$

<sup>2</sup> If relative price parity were to hold  $P_{jkt}^i$  would equal 0.

<sup>3</sup> When a pair of locations ( $j, k$ ) is considered, dummies for location  $j$  and location  $k$  are equal to 1 and the rest location dummies are 0.

Different measures of price volatility (dependent variable) can be used in the model. For instance,  $P_{jkt}^i$  can be defined as  $\log(P_{j,t}^i / P_{k,t}^i)$  (and not first difference of logs as suggested previously), which would reflect percentage difference between the average prices (not the relative) of product  $i$  in jurisdictions  $j$  and  $k$  at time  $t$ <sup>4</sup>. Standard deviation is not the best measure of volatility because it gives too much weight to outliers, so it might be a good idea to consider a spread between the 10<sup>th</sup> and 90<sup>th</sup> percentile in the time series of  $P_{jkt}^i$  or inter-quartile range (75<sup>th</sup>–25<sup>th</sup> percentile) instead. For the sake of comparison, average  $P_{jkt}^i$  over time series can be also used as a dependent variable in regressions, although it is necessary to keep in mind that then outliers would be given even higher weight than in case of standard deviation.

There is usually a problem with heteroscedasticity of error terms in such models. To account for it one should use White's heteroscedasticity consistent standard errors when estimating test statistics [1, 8]. An alternative model specification where all the variables are divided by log of distance can be also introduced, since it is generally believed that the variance of the error terms is greater for more distant locations:

$$V(P_{jkt}^i) / \ln dist_{jk} = \beta_1 + \beta_2 \left( \frac{Border_{jk}}{\ln dist_{jk}} \right) + \sum_{m=1}^N \gamma_m^i (D_m / \ln dist_{jk}) + v_{jk}. \quad (7)$$

In order to have some intuition about the dynamics of the border effect, one can either split the sample into 2 or more subperiods<sup>5</sup>, consider 2 separate years or just use  $P_{jkt}^i$  and not its volatility across time series  $V(P_{jkt}^i)$  as a dependent variable in the basic regression following Parsley and Wei [11]. Regressions for two periods should be run and then the size of border dummies received from these two regressions must be compared. Robustness of the results can be insured through a split of the sample or exclusion of several periods or goods from it. Where to split the sample and which periods and goods to exclude depends on the individual characteristics of the data set under consideration.

We suggested a number of modifications to the standard methodology. First of all, we tried to augment the model with several additional explanatory variables in order to find the influence of dif-

ferent factors on the relative price volatility and disentangle various determinants of the border effect. For example, we introduced wage volatility into the regression to test a hypothesis that labor market segmentation explains a part of the border effect. To control for possible pricing-to-market behavior of the firms we included variability of gross added value per capita in regression because it can be a proxy for the differences of people's wealth across oblasts, which in turn influence consumers' willingness to spend certain amount of money on a particular product. Apart from that, we added some social and political explanatory variables in the regression. For example, introduction of the relative percentage of people who voted for Yushchenko in 2004 (or for pro-Yushchenko parties in 2006) into the model may reveal whether political preferences play a direct role in explaining price dispersion across the East and the West of Ukraine. Use of relative percentage of Russian-speaking people as another explanatory variable would allow to control for the impact of language differences on the price discrepancy between the Eastern and the Western Ukraine. Besides, we tried to use the common administrative border dummy, which takes the value of 1 whenever two locations have a common administrative (oblast) border and 0 otherwise as explicative variable because neighboring oblasts are likely to have less variation of prices. Finally, a large Dnipro river flows through Ukraine dividing it in half, so it seems reasonable to control for possible «river» effect through introduction of respective dummy into the regression.

Within the border effect framework we also examined 205 different possible East-West divisions, which was generated from the map of Ukraine, in order to find «true» economic border from the data. The 205 pooled regressions for each of these borders with correspondent border dummy have been done, and have been found which border has the largest effect on relative price variation<sup>6</sup>.

## Data and Sample

The main data used for the purpose of this research are monthly average retail prices of different consumer products across the oblasts, which allow to compute relative price volatility. They come from official sources, namely statistical collections «Average Prices and Tariffs for Consumer Goods

<sup>4</sup> It would equal 0 if average price parity were to hold.

<sup>5</sup> Price dispersion can be computed for each year, in order to examine the evolution of the border effect.

<sup>6</sup> We look at coefficients of significant border dummies in different pooled regressions and find which of them has the largest size relatively to distance.

and Services» published by the State Committee of Statistics, for the period from 1997 to 2004. They provide monthly average prices of 29 food products, 35 nonfood products and 21 services across 24 oblasts of Ukraine and Autonomous Republic of Crimea. These are actual prices including indirect taxes such as tax on added value (VAT) and excise tax. Price information is collected in oblast and rayon centers, which are chosen taking into account quantity of urban population and satiation of consumer markets with goods and services. In order to compute average prices, price data are weighted on the share of the urban population.

The data set of prices utilized for the purpose of this research has several advantages. First of all, it comes from official sources. Also, it provides average retail prices of consumer goods, and not price indexes. So, there is no aggregation bias there, which is usually present in price index data [3]. In addition, most of products are narrowly defined, which also reduces the possibility of bias. Another advantage of the data using in this research is coverage of the wide range of different products (85 overall). All the investigations that we came across are based on the price data for much smaller number of products [2, 3, 4, 9, 11]. Having average price data for 25 oblasts we can obtain 300 relative prices for each product. When pooling the data over 85 products we receive cross-section data set with 25 500 observations.

On the basis of available price data we compute relative prices for each product over 300 oblast pairs, their logs and first difference of logs. According our finding, relative prices of food products on average are most close to 1, of services - least close and of nonfood products - somewhere in the middle. Most of services are non-tradable, which explains why on average their relative prices diverge the most from 1. However, from the perspective of tradability one would expect absolute PPP to hold the best for nonfood products and, consequently, their prices to be the closest to 1, since food products are perishable goods, which puts some restriction on their tradability. But, on the other hand, nonfood products are much more heterogeneous than food, which can ration higher variability of their prices. It is probably also due to differentiation of nonfood products that the range between their maximum and minimum relative prices is the highest among product categories.

We consider separately percentage differences of the products' absolute and relative prices<sup>7</sup> for

oblast pairs in which both oblasts are located in the East (East-East), in the West (West-West) and for those pairs in which one oblast is in the East and the other is in the West (East-West). If the East-West border did not matter, then average percentage differences of average and relative prices would be the same for the East-West, East-East and West-West pairs of oblasts. In our case, average percentage differences both for absolute and relative prices are the highest for oblast pairs located in the West, and the lowest for oblast pairs located in the East. Cross-border pairs are in the middle. One would expect price volatility to be higher in the West because Eastern and Western regions according to our classification are not symmetric: the West comprises of almost twice as many oblasts as the East (16 and 9 respectively). Therefore, there are 120 oblast pairs in the West and only 36 in the East. It is more difficult to explain why price volatility for the Western oblast pairs is a little higher than for cross-border pairs. In the paper of Engel and Rogers intra-national price volatility between the US states for some categories of products was also higher than US-Canada cross-border volatility [4]. They explained it by high product differentiation of some products and the fact that there are products, which both Canada and the US mostly import from some third countries.

We also constructed 4 hypothetical baskets of consumer products: first basket comprising 29 food products, second - 35 nonfood products, third - 21 services and forth - all 85 products. To do this we first found the average price of each good in each oblast during the period for which price data of this particular product is available. Then, we used rather primitive construction procedure simply giving all the products in each basket equal weights. It obviously does not have to correspond to reality. Still it allows to make a rough judgment about deviations of price levels across the oblasts. Figure 1 illustrates the results for the basket of all 85 products.

The average price of basket consisting 85 products during 1997-2004 was the highest in Crimea, Dnipropetrovska, Donetska, Luganska and Chernivetska oblasts, all but 1 of which are in the East. It was the lowest in Vinnytska, Volynska, Zhytomyrska, Ivano-Frankivska, Kirovogradska, Ternopil'ska, Kharkivska and Chernigivska oblasts, all but 1 of which are in the West. So, price level is generally higher in the Eastern Ukraine and lower in the Western, which is also shown in Figure 2.

<sup>7</sup> Logs of relative prices and first differences of logs of relative prices respectively.

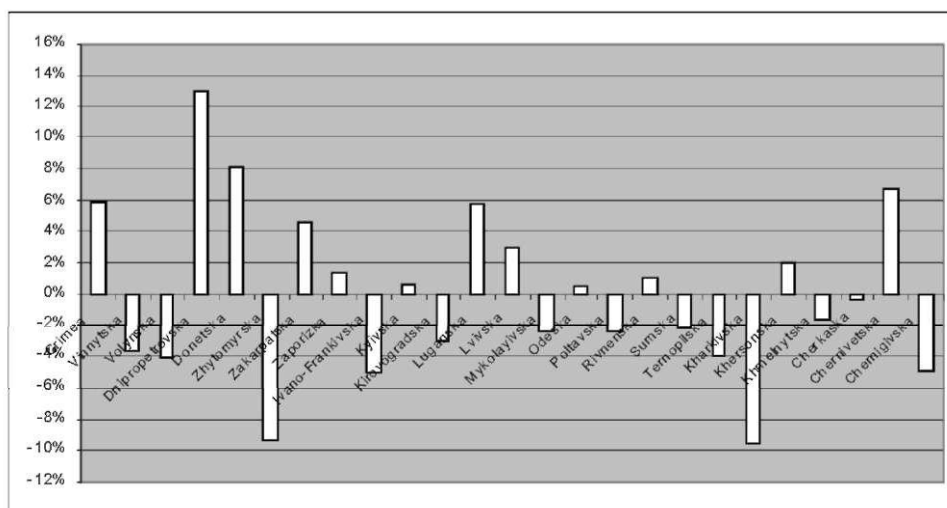


Figure 1. Deviations from the average price of the basket of 85 products across the oblasts of Ukraine

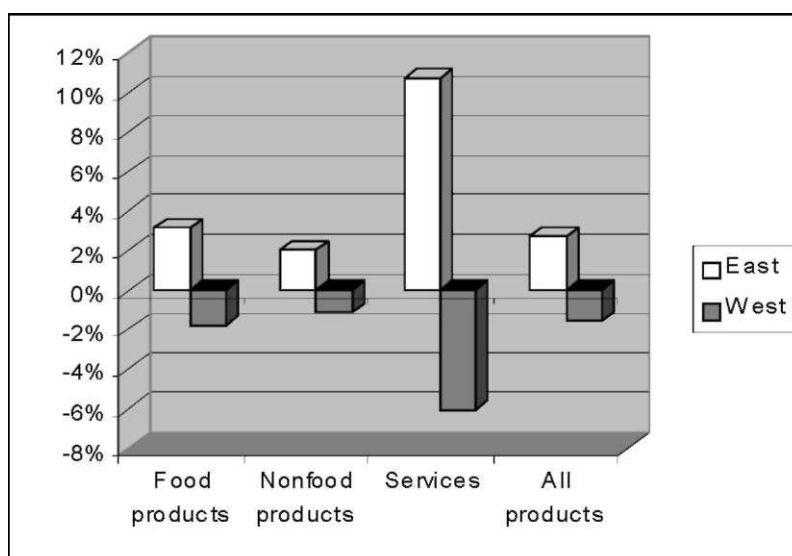


Figure 2. Deviations from the average price of 4 baskets of products in the East and the West of Ukraine

For all 4 baskets of products prices in the East were higher and in the West lower than on average in Ukraine. The biggest difference in price levels between the East and the West is observed for services in line with their non-tradability.

It is also worth to compare average wages across Ukrainian oblasts, since they reflect price of the labor, which is an important factor of production of goods and services. Figure 3 is a counterpart to Figure 1 for wages. It demonstrates deviations from the average level of wages across Ukrainian oblasts

Deviations of wages are an order of magnitude higher than of prices, which is expected because labor is not as mobile as products. Wage deviations range from about -130 % to + 130 %, whereas the spread of price discrepancy is (-10 %; + 13 %).

15 out of 25 oblasts have the same sign of deviations for prices and wages, and among those that have different signs most oblast have deviations rather close to zero. Overall, the highest wages during 1999-2004 were observed in Dnipropetrovska, Zaporizka, Kyivska, Luganska, Mykolayivska and Sumska oblasts, all of which but Kyivska are located in the Eastern Ukraine. The lowest wages had Volynska, Zhytomyrska, Ivano-Frankivska and Lvivska oblasts, all of which are in the Western Ukraine. So, both wages and prices were generally higher in the East than in the West over the period of 1999-2004. To be precise, wages in the East were about 44 % higher and in the West 27 % lower than on average in Ukraine.

The correlation matrix of correlation between average price level<sup>8</sup>, wage, gross added value per

<sup>8</sup>We proxy it to the price of the basket of 85 products that we constructed earlier.

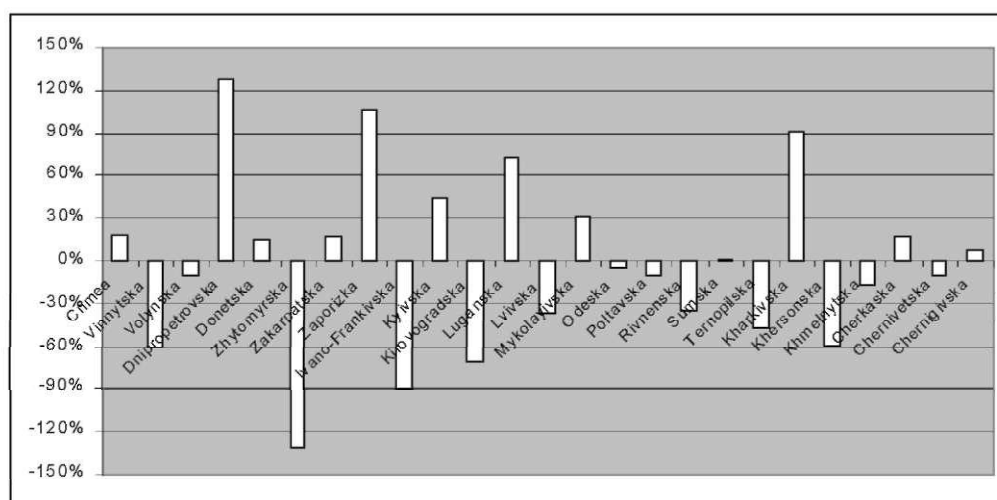


Figure 3. Deviations from the average level of wages across the oblasts of Ukraine

capita, percentage of Russian-speaking people and those who voted for Yanukovich at the Presidential elections 2004 and East dummy<sup>9</sup> for the oblasts of Ukraine was calculated. The main message of this matrix is that there is a substantial positive correlation between all these indicators. They have higher values in the East than in the West, which was earlier shown explicitly for prices and wages. This suggests that if in the empirical part we find a significant border effect between the East and the West it may reflect either economic (wage, gross added value per capita), or political, or linguistic differences, or possibly, some other factors. To check the role of each of these indicators, they will have to be introduced one way or another into the model.

### Empirical results

In the first step a pooled regression (3) from the methodology section 2. has been run, which has exactly the same form as the baseline model offered initially by Engel and Rogers:

$$V(P_{jkt}) = \beta_1 \ln dist_{jk} + \beta_2 Border_{jk} + \sum_{i=1}^K \lambda_i G_i + \sum_{m=1}^{N-1} \gamma_m D_m + u_{jk}, \quad (8)$$

with first difference of logs of relative prices as a dependent variable and log of distance, border dummy<sup>10</sup>, 85 product dummies and 24 oblast dummies<sup>11</sup> as explicative variables (this gives 25 500<sup>12</sup> observations and, therefore, a lot of degrees of freedom, which insures high precision of the estimates). We use White's heteroscedasticity-consistent standard errors, since variance of the error terms is likely to have positive correlation with distance between the locations. Short summary of the results is presented in the table 1.

For this particular regression coefficients of both natural log of distance and border dummy are highly significant. Both of them are greater than zero, which corresponds to theoretical predictions, since one would expect price dispersion to be higher for more distant oblasts and those separated by a border. According to the results of regression, increase of distance between oblasts by 1 % raises price dispersion by 1.76 %. Economic significance of the border can be computed according to the formula proposed by Parsley and Wei [11]:

$$Z = \overline{dist} * \left( \exp\left(\frac{\beta_2}{\beta_1}\right) - 1 \right) = 762 * \exp(0.0030536 / 0.0014601) - 1 = 560 \text{ km.}$$

Table 1. Summary of Stata output for regression (3)

Variable	Coefficient	Robust Std. Error	t	P >  t	95 % Confidence Interval	
Lndist	.0017601	.0006812	2.58	0.010	.0004249	.0030953
Border	.0030536	.0007297	4.18	0.000	.0016233	.0044839

<sup>9</sup> Takes on value 1 whenever oblast is in the East according to the political division, and 0 otherwise.

<sup>10</sup> For East-West political border.

<sup>11</sup> One oblast dummy has to be excluded to avoid perfect collinearity, we excluded dummy for Chernigiv oblast.

<sup>12</sup> 85\*300 (number of oblast pairs) = 25500.

which is negligibly small value in comparison with findings of Engel and Rogers [4], who estimated the effect of Canada-US border to be equivalent to 75000 miles.

R-squared for this regression equals to 0.7724, so the model has rather high explanatory power. Coefficients of all 85 product dummies are highly significant (p-value = 0.000), suggesting that price volatility has some important product-specific features. 21 out of 24 oblast dummies<sup>13</sup> are significant at 5 % level of significance, which means that oblast-specific characteristics also have substantial impact on price dispersion. For instance, some oblasts might have more integrated markets with the rest of Ukraine than the other. Then, price volatility for oblast pairs containing these oblasts would be lower on average, and vice versa. In a given regression oblast pairs that include either Kirovogradska, or Ternopil'ska, or Cherkaska oblasts appeared to have price dispersion above mean.

If to run regression (3) but exclude border dummy from it, R-squared will remain essentially the same but the coefficient of logged distance will be twice as high as in the original regression. This will happen because now the coefficient will show not only the effect of distance but also, implicitly, effect of the border - omitted variable in this specification. Even in the original regression (3) there could be a misspecification bias because historical borders and/or Dnipro River might also matter for the magnitude of price dispersion between the oblasts in Ukraine. Besides, oblasts that share common border might have lower relative price volatility. So, in the second step we will run regression (3) augmented by historical and common border dummies and the Dnipro River dummy. We constructed the Dnipro River dummy the way that it takes on value 1 any time oblast pair contains oblasts located on different sides of the Dnipro River, and 0 if they are located on the same side. However, there are some oblasts (Dnipropetrovska, Kyiv'ska, Kherson'ska and Cherkaska), which are crossed by the Dnipro river in the middle. We as-

Table 2. Summary of Stata output for regression (4)

Variable	Coefficient	Robust Std. Error	t	P >  t	95 % Confidence Interval	
Distance	9.57e-06	4.06e-06	2.36	0.018	1.62e-06	.0000175
Dist-sqr	-5.11e-09	3.06e-09	-1.67	0.095	-1.11e-08	8.97e-10
Border	.0031926	.0007576	4.21	0.000	.0017078	.0046775

sume that in these oblasts the river effect is already incorporated in the intra-oblast price dispersion, so for pairs containing these oblasts the Dnipro River dummy always equals to zero. From a theoretical standpoint, we would expect coefficients of the historical border and Dnipro dummies to be positive, and of common border dummy - negative. These predictions hold for coefficients of Dnipro and common border dummies but not of historical border. But, actually, it is not very important, since all of them are insignificant anyway. So, no evidence that the regression (3) has misspecification bias is found so far. In the next step, running regression (4) with quadratic specification of distance proves concave relationship of distance: coefficient of distance is positive and statistically significant at 5 % level of significance, and coefficient of squared distance - negative and significant at 10 % level. Border dummy coefficient remains positive and highly significant in this specification. Results are present in the table 2.

Controlling explicitly for heteroscedasticity (running regression (7):

$$V(P_{jkt}^i) / \ln dist_{jk} = \beta_1^i + \beta_2^i (Border_{jk} / \ln dist_{jk}) + \sum_{m=1}^N \gamma_m^i (D_m / \ln dist_{jk}) + v_{jk}$$

does not alter general results. Coefficient of the border dummy remains significant and approximately of the same size (0.003412). Since most of 24 oblast dummies are statistically significant in all the mentioned specifications, we find it reasonable to try to include in the model 299 dummies for all but one oblast pairs. The rationale is that if oblast-specific features influence substantially price dispersion, then, possibly, oblast pair-specific features also do. Inclusion of 299 more explicative variables is not going to hurt degrees of freedom too badly because we have very large number of observations - 25 500. But after all we find out that coefficients of only about a dozen out of 299 oblast pair dummies are significant at 5 % level of significance, and just one - at 1 % significance level.

<sup>13</sup> To save space, we do not provide Stata output for 85 product dummies and 24 product dummies. Henceforth only results for the variables of special interest are provided.



In all these steps we use just one measure of price volatility, namely standard deviation of the first difference of logs of relative prices. But it is worthwhile to consider also 7 other price volatility measures mentioned in the section on methodology, namely: standard deviation of  $\log(P_{j,t}^i/P_{k,t}^i)$  (volatility 2); spread between 10<sup>th</sup> and 90<sup>th</sup> percentile of  $P_{jkt}^i$ , where

$$P_{jkt}^i = \log\left(\frac{P_{j,t}^i}{P_{k,t}^i}\right) - \log\left(\frac{P_{j,t-1}^i}{P_{k,t-1}^i}\right)$$

(volatility 3) or  $\log(P_{j,t}^i/P_{k,t}^i)$  (volatility 4); interquartile range of  $P_{jkt}^i$  (volatility 5 and 6); mean  $P_{jkt}^i$  (volatility 7 and 8). We duplicated main points of our analysis for these volatility measures. Short summary of results can be found in the table 3.

As noted in section 2, measures of volatility 3–6 ignore outliers, whereas volatilities 7-8 give them rather high weigh. According to Table 3, for all the volatilities ignoring outliers R-squared is rather

high, log of distance has big explanatory power, whereas political border appears to be insignificant. For volatilities which take outliers into account the situation is the opposite. Quite logically, R-squared is very small for them because outliers usually reflect some shocks. However, it is a bit surprising that logged distance has no substantial impact on them, whereas border is important.

For regressions with volatilities 2-6 as dependent variables the Dnipro River, historical and common border dummies were insignificant. However, for volatilities 7-8 common border dummy becomes highly significant (p-value = 0.001 and 0.000 respectively) and has negative sign as expected, since oblasts that have a common border are supposed to have more integrated markets and, therefore, lower price dispersion. Roughly 95 % of all product dummies and 75 % of oblast dummies are significant for all these volatility measures, so product-specific and oblast-specific effects repeatedly prove to be important.

Table 3. Main Stata output for specification (3) with volatility 2 to 8 as a dependent variable

Variable	Coefficient	Robust Std. Error	t	P >  t	95 % Confidence Interval	
<b>Volatility 2</b> (R-sqr = 0.8259)						
Lndist	.0053094	.0007702	6.89	0.000	.0037998	.0068191
Border	.0010746	.0008894	1.21	0.227	-.0006686	.0028178
<b>Volatility 3</b> (R-sqr = 0.8938)						
Lndist	.003103	.0006455	4.81	0.000	.0018377	.0043682
Border	.0003144	.0007991	0.39	0.694	-.0012519	.0018807
<b>Volatility 4</b> (R-sqr = 0.8315)						
Lndist	.0150636	.0018978	7.94	0.000	.0113438	.0187834
Border	.001789	.0022003	0.81	0.416	-.0025238	.0061018
<b>Volatility 5</b> (R-sqr = 0.9074)						
Lndist	.0014993	.0002471	6.07	0.000	.0010151	.0019835
Border	.0000591	.000304	0.19	0.846	-.0005366	.0006549
<b>Volatility 6</b> (R-sqr = 0.7646)						
Lndist	.008711	.0012356	7.05	0.000	.0062892	.0111327
Border	.0013757	.0015331	0.9	0.370	-.0016294	.0043807
<b>Volatility 7</b> (R-sqr = 0.0485)						
Lndist	.0000294	.0001804	0.16	0.871	-.0003242	.0003829
Border	.0005355	.0002217	2.41	0.016	.0001008	.0009701
<b>Volatility 8</b> (R-sqr = 0.0821)						
Lndist	-.0032037	.0031969	-1	0.316	-.0094698	.0030624
Border	.0208572	.0039093	5.34	0.000	.0131948	.0285196

The analysis of different volatilities has already shown that the results obtained from regression (3) are not robust, so there is not much sense to present any other check on robustness like split of the sample or exclusion of some products. We proceed further by testing the importance of wage volatility, differences in the average gross added value per capita, percentage of people who voted for Yushchenko in 2004 and share of Russian-speaking population in explaining price dispersion across oblasts.

For the analysis of the impact of wage volatility on price dispersion we use two different ways to compute wage volatility, which are essentially counterparts to the first and second measures of price volatility: standard deviation of  $w_{jkt}^i$  with

$$w_{jkt}^i = \log\left(\frac{w_{j,t}^i}{w_{k,t}^i}\right) - \log\left(\frac{w_{j,t-1}^i}{w_{k,t-1}^i}\right)$$

(wage volatility 1) or  $\log(w_{j,t}^i/w_{k,t}^i)$  (wage volatility 2).

Since there are differences in periods for which price data for various products are available, as stated in section 3, we compute separately measures of wage volatility correspondent to each product and then pool them over all the products. However, some disparities can not be eliminated because price data for 30 out of 85 products are available starting from 1997, whereas wage data are available only from 1999. After computing wage volatilities we run two pooled regression of the general form:

$$V(P_{jkt}) = \beta_1 \ln dist_{jk} + \beta_2 Border_{jk} + \beta_3 V(w_{jkt}) + \sum_{i=1}^K \lambda_i G_i + \sum_{m=1}^{N-1} \gamma_m D_m + u_{jkt}. \quad (10)$$

In pooled regression

$$\text{with } P_{jkt}^i = \log\left(\frac{P_{j,t}^i}{P_{k,t}^i}\right) - \log\left(\frac{P_{j,t-1}^i}{P_{k,t-1}^i}\right)$$

$$\text{and } w_{jkt}^i = \log\left(\frac{w_{j,t}^i}{w_{k,t}^i}\right) - \log\left(\frac{w_{j,t-1}^i}{w_{k,t-1}^i}\right)$$

coefficient of wage volatility is statistically significant at 5 % level of significance but surprisingly has a negative sign. These unexpected results could be partly due to lack of correspondence between

the periods for which price volatility and wage volatility are computed as mentioned above.

To test the significance of differences in political and linguistic preferences, we use data on percentage of people who voted for Yushchenko at Presidential elections 2004 and percentage of people who consider Russian their native language to compute differences for all oblast pairs. This allows us to receive two explicative variables - proxies of political and linguistic preferences. We add them to model (3) and run 8 pooled regressions with different measures of price volatility.

Differences in political preferences appear to have no direct impact on the price dispersion, since coefficient of this variable is insignificant in all 8 specifications. A possible explanation can be that political preferences in reality are important but have to enter regression in a different functional form.

There is some evidence, however, about positive influence of differences in linguistic preferences on price dispersion. Its coefficient is positive and statistically significant for specifications with volatilities 6, 7 and 8. The results of the estimation is shown in the table 4.

This corresponds to theoretical predictions and means that the more different are two oblasts according to language preferences, the higher price dispersion one might expect for them. Besides, native language might reflect person's origin and, therefore, some cultural differences, including preferences what products to consume.

Analysis of variation in gross added value per capita across oblasts, which can be used as a proxy for income and wealth, also produces some interesting results reproduced in the table 5.

This variable appeared to be significant in 4 out of 8 pooled regressions and always higher than zero. This means the more different are two oblasts in terms of wealth the higher price dispersion can be expected for the them. It can also potentially mean pricing-to-market behavior of the firms.

We also run simple OLS regressions for each of 85 products separately with volatility 1 and volatility 2 as dependent variables. The separate OLS regressions for 85 products do not give strong evidence on the border effect. Running pooled regressions with standard deviation of  $P_{jkt}^i$  as a de-

Table 4. Testing for significance of linguistic differences

Variable	Coefficient	Robust Std. Error	t	P >  t
Language (volatility 6)	.0170623	.0013244	2.05	0.040
Language (volatility 7)	.0033063	.0012093	2.73	0.006
Language (volatility 8)	.0796957	.021798	3.66	0.000

Table 5. Testing for significance of the gross added value (GVA) per capita

Variable	Coefficient	Robust Std. Error	t	P >  t
Relative GAV per capita (volatility 1)	.002171	.0012135	1.79	0.074
Relative GAV per capita (volatility 3)	.0023648	.0010939	2.16	0.031
Relative GAV per capita (volatility 7)	.0056722	.0003312	17.13	17.13
Relative GAV per capita (volatility 8)	.1711945	.0058568	29.23	0.000

pendent variable

$$\left( P_{jkt}^i = \log\left(\frac{P_{j,t}^i}{P_{k,t}^i}\right) - \log\left(\frac{P_{j,t-1}^i}{P_{k,t-1}^i}\right) \right)$$

and  $P_{jkt}^i = \log(P_{j,t}^i / P_{k,t}^i)$  allows drawing the following conclusions: distance is highly significant in both specifications; political border is highly significant in the first specification but not in the second; historical border is statistically insignificant in both specifications<sup>14</sup>; common border dummies are statistically insignificant in both specifications; around 95 % of oblast dummies are highly significant in both specifications; around 80% of product dummies are highly significant in both specifications; river dummies are statistically insignificant in both specifications. Adjusted  $R^2$  equals about 0.7723 in specification1 and about 0.8259 in specification 2. Introduction of the border dummy into regression does not increase adjusted  $R^2$  substantially but reduces the size of the distance coefficient.

We also tried running pooled regressions with quadratic specification of distance. This does not provide strong evidence for the border effect in Ukraine, since coefficient of the East-West border is insignificant in most of separate OLS regressions for 85 products. However, it demonstrates that the effect of this border on the price dispersion is larger than of historical border, common borders and Dnipro River.

We also tried finding the 'true' East-West border suggested by data. In order to do this, we generated 205 potentially possible East-West divisions from the map of Ukraine. Then, we created correspondent border dummies, and ran regression (3) for 8 different measures of price volatility and 205 different East-West border dummies (1640 regressions in total). We found borders, which were significant and had the highest values for most of these 8 price volatilities. According to this analysis, the

political border was among the best candidates for the 'true' East-West border.

### Conclusions

Our analysis is the first to provide the investigation of the border effect among different regions of Ukraine as well as between the East and the West parts of our country. We were first to answer the very important question: how far the West is from the East along different dimensions (economic, political, cultural and others), whether markets are segmented according to voting patterns, and whether there exists an objective risk of the division of Ukraine into two parts.

The empirical study showed that there is some evidence that political border has a positive impact on the price dispersion across the Ukrainian regions but its economic significance is not very high. When converted into distance units this border is equivalent to about 560 kilometers, which is rather low figure in comparison with findings of other researchers for other countries. Not all of the measures of price volatility reveal positive and significant border effect, so the results are not robust. Still, the political border appears to have much higher explanatory power of the price dispersion than historical border, common borders between oblasts, the Dnipro River and wage volatility. Moreover, political border is a good candidate for the 'true' East-West border of Ukraine according to actual data. So, its role definitely should not be ignored.

However, some fixed product-specific and oblast-specific features explain much larger part of the price volatility, which suggests that Ukrainian markets are more segmented by product and oblast than by hypothetical «East-West» border. Distances between locations, which approximate well shipping costs, are also consistently proven to have a positive effect on the price dispersion. Also, differences in the linguistic preferences and gross

<sup>14</sup> Inclusion both political and historical border dummies into regression simultaneously does not change the results.

added value per capita appear to matter for price volatility across the oblasts of Ukraine. The impact of the former could reflect people's heterogeneity by origin, which influences their consumption preferences. Both are likely to be connected to the pricing-to-market behavior of the firms, which implies that Ukrainian markets are not very competitive. These findings can have some important policy implications for Ukrainian government: language issues should be given more attention, since they

matter not only from the social and political perspectives but also have some economic meaning; economic competition policies must be reviewed and improved to insure the competitive environment and lack of opportunities for pricing-to-market behavior of the firms. However, the evidence of economic nature of the East-West political border does not seem strong enough to support the need of transforming Ukraine into the federal state.

1. Лук'яненко І. Г., Городніченко Ю. О. Сучасні економетричні методи у фінансах. - К.: Літера ЛТД, 2002. - 352 с.
2. Anderson M., Smith S. Do National Borders Really Matter? A Reconsideration of Canada - US // Regional Trade Review of international economics. - 1999. - № 7. - P. 219-227.
3. Ceglowski J. Law of One Price: International Evidence for Canada // The Canadian Journal of Economics. - 2003. - 36(2). - P. 373-400.
4. Engel C., Rogers J. How Wide Is the Border? // The American Economic Review. - 1996. - 86(5). - P. 1112-1125.
5. Engel C., Rogers J. Violating the Law of One Price: Should We Make a Federal Case Out of It? // Journal of Money, Credit and Banking. - 2001. - 33(1). - P. 1-15.
6. Frankel J., Rose A. A Panel Project on Purchasing Power Parity: Mean Reversion within and between Countries // The Journal of International Economics. - 1996. - 40(1). - P. 209-224.
7. Gorodnichenko Y., Tesar L. A Re-Examination of the Border Effect // NBER Working Papers 11706. - 2005.
8. Greene, William. Econometric analysis. - 4<sup>th</sup> ed. - Upper Saddle River, New Jersey: Prentice-Hall, Inc., 2000. - 1004 p.
9. Helliwell J. Do National Borders matter for Quebec's Trade? // The Canadian Journal of Economics. - 1996. - 29. - P. 507-522.
10. McCallum J. National Borders Matter: Canada - US. Regional Trade Patterns // American Economic Review. - 1995. - 85. - P. 615-623.
11. Parsley D., Wei S. Explaining the Border Effect: The Role of Exchange Rate Variability, Shipping Costs, and Geography // The Journal of International Economics. - 2001. - 55(1). - P. 87-106.
12. Rogoff, K. The Purchasing Power Puzzle // The Journal of Economic Literature. - 1996. - 34. - P. 647-688.
13. Wolf, H. Intranational Home Bias in Trade // The Review of Economics and Statistics. - 2000. - Vol. 82(4), P. 555-563.

Лук'яненко І. Г., Городніченко Ю. О., Гузенко С. І.

## ПРОБЛЕМИ ОЦІНКИ ЕФЕКТУ КОРДОНІВ В УКРАЇНІ

*Пропонована стаття - одна з перших спроб емпіричного дослідження існування уявного кордону між Сходом та Заходом України. Було доведено, що ефект кордону незначний, а відстань, яка йому відповідає, не перевищує 560 км, що набагато менше порівняно з існуванням адміністративних кордонів в інших країнах світу. Нами було також виявлено основні чинники, що визначають межі в Україні. Дослідження, проведене на основі економетричного інструментарію, надало можливість зробити загальний висновок щодо відсутності явного кордону між Сходом та Заходом України, а відповідно, і відсутності об'єктивних причин до трансформації існуючих кордонів та розколу України.*