44TH EUROPEAN CONGRESS OF THE EUROPEAN REGIONAL SCIENCE ASSOCIATION

25-29 AUGUST, 2004, PORTO, PORTUGAL

Economic Forecasting for Large Russian Cities

by

Raymond Struyk, Senior Fellow - The Urban Institute, Frankfurt, Germany,
Douglas Wissoker, expert - The Urban Institute, Washington, DC, USA,
Ioulia Zaitseva, PhD, project manager, The Institute of Urban Economics, Moscow, Russian Federation

The Budget Code of the Russian Federation requires that local self-governments prepare their budgets for the next year taking into account the likely economic situation in that year. To date these governments have had little guidance to use in preparing their budgets. This paper reports the results of initial steps to develop a procedure for forecasting key economic parameters at the local level. "Local level" is defined as cities that are capitals of Subjects of the Federation (similar to U.S. states); generally these are cities of over 100,000 population. Econometric models are reported for employment, manufacturing production, retail sales, average wage rates, volume of newly constructed housing, and fixed capital formation. The choice of estimation procedures was significantly constrained by data availability. The current document is an interim report, prepared after the basic econometric work has been completed but before the model is tested in actual forecasting. The paper consists of six further sections. The first lists the economic variables to be projected. The second describes the economic logic underlying the models specified for each variable. The third section then outlines the econometric strategy. This is followed in the fourth section with an overview of the data employed in the estimates. The fifth section presents the final models. The paper closes with a short discussion of the plans for future work in this direction. In the next phase of the work the forecasting qualities of these models will be evaluated.

Economic Variables of Interest

Analysts at the Institute for Urban Economics with extensive experience working with cities on economic development and local government finance issues determined the variables of greatest interest. However, their deliberations were limited by the variables for which the necessary data are available. The list of variables to be projected by the analysis, finally determined through this iterative process, is shown in the first column of the following table.

Table 1. List of Economic Variables to be Forecast

Variable	Label	Operational Definition
Employment	EMPPC	Employment per 1,000 population
Manufacturing production	MFGPC	Value of manufacturing output per capita in constant prices[a]
Retail sales	RETAIL-PC	Value of retail sales per capita in constant prices
Wage rate	WAGE-LVL	Average level of wages and salaries in constant prices
New housing constructed	NEWHSG	Square meters of new housing commissioned per capita
Fixed capital formation	FIXCAP-PC	Value of fixed capital put in place per capita in constant prices

a. 2001 is the base year for the price deflator.

The second column shows the label of the variable used elsewhere in the report and the third provides the operational definition ultimately adopted. Levels rather than growth rates are employed to incorporate into the prediction the long-term relationship between the levels. Most of the variables are standardized by defining them on a per capita basis to control partially for the large differences in the size of the local economies included in the analysis.

The Economic Logic

The general procedure for developing the forecasts of the economic variables was to begin by specifying and estimating a regression model for each of these variables. The resultant equations are used in making the forecasts.

This section discusses the variables included in each of the estimated regression models, i.e., the underlying economic logic for each independent variable in an equation.

Broadly, there are three types of independent variables employed in the analysis that we somewhat arbitrarily label as causal variables, lagged outcomes, and control variables. The causal variables are those that have a clearly defined economic relationship with the dependent variable. For example, in a model explaining the level of employment per capita in a given year, the previous year's level of regional GDP is a causal variable—a positive or negative

change in GDP will produce a corresponding change in the employment level. These are highly dynamic factors and we have annual observations on their values. The lagged values are included for their predictive power. Control variables with one exception are indicators of region, climate, and city size that control for relatively stable, broad differences not captured by the causal and lagged variables. A further control variable is for the negative economic jolt caused by the 1998 economic crisis.

The structure of each causal model was based in part on the experience of other researchers estimating regional forecasting models. We also examined the literature on analysis of interregional variations in wage rates, investment, unemployment levels and other factors. Finally, recent contributions to the literature on the economics of Russian regions were consulted.[i]

Based on this information, we specified the trial model specifications shown in Table 2. The logic for each model is summarized below.

EMPCC—employment per capita. Employment is driven by the level of economic activity in the area and is impeded by high wage rates, which discourage additional investment being made. Two specifications for regional economic activity were tried: total regional GDP and then the major components of GDP.

MFGPC—manufacturing output per capita. Manufacturing depends on the level of economic activity and is reduced by high wage rates, which make the city less competitive and discourage additional investment. Depending on the type of product manufactured, much of the production could be shipped outside of the region and therefore the level of manufacturing output could depend at least as much on the national economy. (This point is discussed further below.)

RETAIL-PC—retail trade per capita. Trade depends in on the level of effective demand. In this specification, demand is captured by the level of regional GDP and the local wage rate, with higher wage rates generating greater purchasing power, other factors being the same.

WAGE-LVL—average level of real wages and salaries. Wages are driven by the demand for labor, captured here by regional GDP, the price level in the area, and by housing space per person. Given the wide variation in living costs among Russia's regions, variation in the CPI would be expected to an important determinant of nominal wages. But the relationship between real wages and regional CPI over time is less clear and we have no strong hypothesis about its direction. Regarding housing space, studies of labor mobility and wage rates in Eastern Europe and CIS countries have found that workers demand a wage premium to live in areas with housing shortages.[ii]

NEWHSG—square meters of new housing commissioned per capita. This is a reduced form equation containing both elements of demand and supply. In one specification, demand depends on the wage level, employment level, the share of the capital budget accounted for by the local administration, which tend to favor housing construction, and the housing space per capita—places with less housing, other factors held constant, will have a greater demand for additional housing space. In the second specification, local wages and employment are replaced by regional GDP. Given the long construction periods, the level of these factors in the previous period are likely to be more important than their levels in the current period. On the supply side, the cost of funds will be important, which is indicated in both specifications by the national loan interest rate.

FIXCAP-PC—the level of fixed capital investment. Investments for the creation and rehabilitation of fixed assets, including the new construction, expansion, and reconstruction and modernization of objects; purchase of machines, equipment, vehicles; and, long-term agricultural investments. The demand for fixed capital is a derived demand, based on the demand for basic goods and services. The two model specifications try to capture the basic demand in two ways: one specification uses regional GDP and the other the level of manufacturing activity. Manufacturing output, with its greater capital intensity compared with other sectors, may be more important than the overall level of demand captured in GDP. Two other factors are included in both models. The first is interest rates, with less investment hypothesized to occur when interest rates are relatively high. The second is the importance of budget funds in total fixed investment, to capture the exogenous demand for fixed capital.

Table 2. Trial "Simple" Model Specifications

	EMPPC	EMPPC	MFGPC	RETAIL-PC	WAGE-LVL	NEWHSG	NEWHSG	FIXCAP-PC	FIXCAP-PC
GDPREG	X		X	X				X	
WAGE-LVL-1	X		X	X		X			
MFGPC-1		X							X
SRVPC-1		X							
RETAIL-PC-1		X							
FIXCAP-PC-1		X							
EMPPC-1					X	X			
CPIREG-1					X				
GDPREG-1							X		
CAPBUG-1						X	X	X	X
HSGPC					X	X	X		
INTEREST						X	X	X	X

Additional variable definitions

GDPREG	Per capita regional GDP (1000s), in constant prices
SRVPC	Value of services per capita in constant prices
CPIREG	Regional consumer price index
CAPBUG	Fixed capital investment from budget sources in constant prices
HSGPC	Housing space per capita
INTEREST	Interest rate on loans of one-year or more

The requirement for control variables arises primarily from using observations from a group of cities to create a set of observations to estimate the econometric models. The control variables account for comparatively static differences in certain conditions across locations in Russia. These include differences in climate and economic structure, for example. As suggested, in the models they help the analyst to control for relatively stable, broad regional differences that are not captured by the causal variables. Variables of this type are listed in Table 3. Those on industrial structure and climate control for factors that could clearly affect the level of economic development. The variables for the Federal District in which the city is located are to account for other sources of variation not associated with climate or industry mix. Variables from this set were ultimately included on a test basis in all of the models.[iii] We refer to these as the "complex models," as opposed to the "simple models" that only contain the causal variables.

The final control variable is for the 1998 economic crisis. The sharp but short-lived economic downturn that resulted in lower economic output, consumption, and wages in that year.

Table 3. Labels and Definitions of Control Variables

Label	Definition
	Region
Central	=1, if city is in the Central Federal District
Northwest	=1, if city is in the Northwest Federal District
South	=1, if city is in the South Federal District
Volga	=1, if city is in the Volga Federal District
Urals	=1, if city is in the Urals Federal District
Siberia	=1, if city is in the Siberia Federal District
Far East[a]	=1, if city is in the Far East Federal District
	Dominant industry[b]
Electricity	=1, if industry accounts for more than 30% total industrial production
Fuel	=1, if industry accounts for more than 30% total industrial production
Metals	=1, if industry accounts for more than 30% total industrial production
Chemicals	=1, if industry accounts for more than 30% total industrial production
Machine-building	=1, if industry accounts for more than 30% total industrial production
Timber	=1, if industry accounts for more than 30% total industrial production
Building materials	=1, if industry accounts for more than 30% total industrial production
Light manufacture	=1, if industry accounts for more than 30% total industrial production
Food	=1, if industry accounts for more than 30% total industrial production
Other-industry	=1, if industry accounts for more than 30% total industrial production
Mixed	=1, if more than one industry each accounts for more than 30% total industrial production
	Temperature – alternative specifications
TEMPJAN	Average temperature in January
TEMPJLY	Average temperature in July
TEMPAVRG	Average temperature over the year
	Size of place[c]
LRGCTY	=1, if city population over 1 million
MEDCTY	=1, if city population between .5 and 1 million
SMLCTY	=1, if city population under 0.1 million
	1998 economic crisis
D1998	=1, for observations in 1998

a. In reported regression models, this is the omitted category.
b. Data for 1998. Some cities have no industry that accounts for 30% of total output; so there is no omitted dummy variable category
c. Omitted category is places between 100,000 and 500,000 population.

Estimation Strategy

The strategy for estimating the models was influenced by the uncertainty we faced in part because of the modest experience with this type of econometric forecasting model for Russia (that we could find to draw upon) and by certain data constraints. There are three aspects to the estimation strategy that warrant comment.

Sequencing. As suggested earlier, we began by estimating single equation models relying solely on causal variables. To test the validity of the causal models specified, we estimated these first, i.e., without including other variables. As the second step, we added auto-regressive structures, i.e., lagged values of the dependent variables. These variables typically added a lot of explanatory power to the models, because a municipality's economic performance next year depends substantially on how it performed in the previous year. Furthermore, the lagged values tend to capture the variation in an outcome across areas. We also experimented with two years of lags to test if there were a longer systematic relation between past events and the current developments than captured in a single year. As the last step, we added the control variables described above. Below, we report the results of these final models, together with an indicator of the contribution of variables other than the lag to the prediction.

Ordinary least squares are used to produce the single-equation estimates. More precise estimates could be obtained using generalized least squares to account for the measurement of the outcome in per capita terms and the resulting inverse relationship between the residual variance and city size. This inverse relationship results when the outcome (e.g., employment per capita) is a sample mean, so its variance–-defined as the population variance divided by the sample size–-is smaller in cities with more persons and larger in cities fewer persons. The generalized least squares procedure would take this heteroskedasticity into account when estimating the regression parameters. Ignoring such a relationship is known to lead to incorrect standard errors and to somewhat less precisely-estimated parameters. Despite this, we chose not to weight. Weighting gives the most emphasis to largest cities and we want to ensure that we can predict as well for small capitals as for large ones.[iv]

We estimated robust standard errors for the OLS regression estimates using a clustered robust variance formula. The resulting standard errors allow variation in the residual variance across observations, as we would expect with heteroskedasticity, and allow correlation of the residuals of observations over time within a given capital. The adjustment to standard errors varies across coefficients and models, with the average increase in standard errors being approximately 50 percent.[v]

Simultaneity. We tested a three-stage least squares, simultaneous equations model in which contemporaneous regional GDP was treated as endogenous, while lagged outcomes, including lagged GDP, were assumed exogenous.[vi] In single equation models, the coefficients on current and lagged GDP were opposite in sign, with a relatively small difference in their magnitudes. In the simultaneous model, the coefficients on the current and lagged values of GDP increased greatly in absolute terms, with little change in their sum. This suggested that

inclusion of both measures was unwarranted. As a result, we decided to drop current GDP from the model and rely instead on the lagged value of GDP for our estimates.

Final specification. Our final model specification involves estimating each equation separately, treating all independent variables as exogenous. The specification includes lagged values for the specific outcome, lagged values of regional GDP and, and a subset of the fixed control variables. None of the variables are treated as simultaneously determined, since all potentially endogenous variables enter with a lag.[vii] As noted above, the tests of the precision of the coefficients are based on robust standard errors.

Data restrictions. There are three restrictions imposed by the data available. First, the conventional approach to estimating regional economic models is to employ time-series data for a single region. In Russia this is not possible because the time series are too short in two senses. One is because the models being estimated are based on market concepts, our observation period must be restricted to those years when Russia's economy was fairly clearly following market principles. The other reason the observation period is restricted is that some analysts have questioned the quality of some data series during the early years of the transition. These considerations caused us to use combined time-series cross-section data in order to have a sufficient number of observations and to limit the data series to years no earlier than 1997. The use of the cross-sectional data required the inclusion of the control variables described earlier. The sample size for most regressions is 320—80 cities for four years.

The second limitation concerns the timing of the availability of the data for each subsequent year. There is a 15-month delay from the end of a year and the time at when the data for that year are released. For example, the data for 2001 were released in April 2003. This could make forecasting perilous because those making the projections will have to project three years forward. Consider the following example. Projections to be used in preparing a local government's 2004 budget would have to be given to the government in November 2003. To prepare the estimates, the analysts would have data for 2001 as the most recent; so they would have to project values for 2002, use the projected values to predict 2003 values, and then use the 2003 predicted values to project the 2004 estimates. Obviously, the more years that one must forecast forward, the less reliable the projections will be.

To minimize the number of variables that we must project three years, we tested the use of one-year lagged values of many of the independent variables. Where these proved serviceable, they are used so that these values only must be projected two years.[viii]

Note, however, that projections of GDP at the regional level are prepared by the Ministry of Economy and Trade. So, current values of this variable could be used. As noted, we found that one-year lagged values of regional GDP performed as well as contemporaneous

regional GDP in the models; so in the results reported below those for lagged GDP are presented. Given the strong dependence of regional GDP on national (but not the reverse), we elect not to use national GDP in the models.

The third data limitation is that data are available only for cities of 100,000 or more population that are capitals of Subjects of the Federation. This limits the number of observations and, more importantly, the size of place for which the projections are likely to be valid.

Data Sources

The main data source was the yearbook issued by Goskomstat, *Regions of Russia*, for the years 2001 and 2002. Most variables come from volume 1, although the regional deflators are in volume 2. Interest rates on loans were taken from the web-site of the Central Bank of Russia (www.cbr.ru). Temperatures were found in the encyclopedia, "Cities of Russia." Other variables were found on the web sites of Goskomstat (www.gks.ru) and the Ministry of Economic Development and Trade (www.economy.gov.ru).

It is important to note that the time series data for individual variables for specific cities contain large year-to-year discontinuities that are difficult to relate to broad economic events. In other words, the data are quite rough, a fact that made it uncertain whether significant and plausible models could be estimated.

Results

The best-performing regression equations are presented in Table 4. Performance was based on the consistency of the signs of the independent variables with our expectations based on broad economic theory for each model, as outlined earlier, the statistical significance of the variables, and the explanatory power of the model. After some general comments, the discussion is organized by dependent variable.

Overall these models performed as expected, although there are some notable exceptions that are described below. All the models are highly statistically significant. Auto-regressive processes were important in all models. As a measure of the degree of importance of other independent variables in the models, the final row in the table records the R^2 of the models in which the first difference of the outcome is regressed on all variables other than lags of the outcome. The R^2 range from .156 to .761— indicating that these variables add a non-negligible amount to a model based only on lagged dependent variables. For several outcomes, the control variables, in addition to that for 1998, have significant effects, indicating that they may be

helpful in forecasting these outcomes. Finally, the significance of most of the control variables proved sensitive to the exact specification of the models.

EMPCC—employment per 1,000 population. Our expectation was confirmed that employment depends on regional per capita GDP. Both specifications for GDP were significant but yielded the same explanatory power; so we elected the simpler specification, i.e., only the summary GDP measure. Local wages were found to be negatively related to employment levels, consistent with the idea that businesses are attracted to low-wage regions, holding other factors constant. EMPCC-1 was highly significant and has a coefficient of .94, indicating a strong auto-regressive process. The 1998 dummy variable was significant and of the expected negative sign. Only one other control variable proved significant—the mean temperature in July: the higher the temperature, the lower the value of EMPCC. This is consistent with comparatively few non- working individuals in cold regions, a well-know phenomenon in Russia's Far North.

MFGPC—manufacturing output per capita in constant prices. Regional GDP per capita turned out not to be significant; its sign was consistently negative, clearly contrary to expectations. So it was dropped from the model. The only significant variable other than the regional dummies is the lagged value of MFGPC itself. Among the control variables, the regional variables proved to be significant. Compared to the Far East Region, all other regions had significantly greater manufacturing output after controlling for the other variables. Manufacturing output was hard hit by the events of 1998, as indicated by the large coefficient and its high level of significance.

RETAIL-PC—retail sales per capita. The sample size here is smaller than for the other variables because information on this variable has only been more recently published by Goskomstat. The lagged value of regional GDP is only a marginally significant determinant of RETAIL-PC. The lagged value of the dependent variable is highly significant but its coefficient is .91 suggesting the importance of other factors in determining the actual value of retail sales for a city. Additional control variables were generally insignificant or very sensitive to the model specification. In the end, besides the 1998 variable, only the average temperature proved consistently significant and fairly robust. The result indicates that retail spending declines as temperatures fall, possibly reflecting fewer shopping opportunities in more isolated, cold-climate communities.

WAGE-LVL—average level of wages and salaries. As anticipated, there is a strong auto-regressive structure in local wages, with the lagged value of WAGE-LVL being highly significant. Increases in the regional CPI are negatively associated with real payment levels, a result for which we have no ready explanation. Regional GDP lagged one period is also

strongly significant, suggesting that it takes some time before positive local growth translated into greater wage demands. The 1998 dummy variable is highly significant confirming that wages plunged after the economic crisis.

We had hypothesized that a relative shortage of housing would lead to higher wages—as workers demand higher payments to live in cramped conditions. In fact, in several of the estimated models, the opposite result was found, with higher payments associated with more square meters of housing. In effect, the causality is reversed from that specified in the model, i.e., higher wages permit households to purchase more housing. In the final specification, however, the variable was insignificant.

The control variables show strong regional effects. Holding other factors constant, all regions except the Northwest demonstrate lower payments relative to the Far East Region. But among these regions only the result for the Urals remained significant in the final specification.

NEWHSG—square meters per capita of new housing commissioned. As hypothesized, the larger the share of public spending in total capital investment, the greater the volume of new housing construction—regional administrations often favor housing over many other types of fixed investment. The volume of housing commissioned in the previous year (NEWHSG-1) is a significant determinant but not that for one year earlier (NEWHSG-2). Also, the variable on the status of the housing situation, square meters of housing per capita, is not significant, although of the correct sign. This may result from two offsetting tendencies. On the one hand, cities with a relative housing shortage may be building more new housing relative to similar cities with better housing conditions; on the other hand, well-to-do cities, are likely to enjoy high construction rates financed by both public and private sources. The 1998 dummy variables is not significant here, presumably because of the long lead time for new housing completions.

We experimented with two other causal variables: interest rates and city size. While these variables were significant before the adjustment of standard errors for clustering effects, in the final model they are not significant.

FIXCAP-PC—real fixed capital formation per capita. The distinguishing aspect of the model is the large magnitude of the coefficient of the lagged dependent variable at 1.30, indicating a substantial reinforcing effect in real terms in fixed investment per capita after controlling for other factors. On its face, this pattern appears hard to sustain and the result may be specific to the estimation period.

Regional GDP per capita is only a marginally significant, positive determinant. On the other hand, total fixed investment from local budget reduces total investment—but only by about half its amount. The result suggests that the administrations that are aggressively investing are encouraging private entities to do so as well. In several specifications including

control variables for size of city showed larger fixed investment for smaller cities as compared with cities of 100,000 – 500,000 persons. In the final model, all city sizes are statistically insignificant. Overall, this equation is the least satisfactory of all in terms of its explanatory power.

The First Testing

The task of the work summarized here consisted of testing the utility of the models for making projections at the municipal level. The plan was that in the fall of 2003 values for 2004 of economic parameters would be projected for up to five cities recruited by IUE. The projections would be based in part on local data for 2002 and 2003 provided by the cities.

Cities participating in testing of model.

The letter inquiring about possible cooperation in the testing of the IUE forecasting model was sent in 12 cities that included all seven federal districts of the Russian Federation. Positive responses and statistical data, including data for 2003, were received from 9 cities:

- Tula, Yaroslavl (Central Federal District)
- Kaliningrad, Cherepovets (Northwest Federal District)
- Engels, Penza (Volga Federal District)
- Tumen (Ural Federal District)
- Tomsk (Siberia Federal District)
- Magadan (Far East District)

Results of calculation and expert adjustments

The predicted values for five economic variables were obtained using equations represented in the table 1 and the data provided by cities. They are shown in Table 5. Analysis of these results was made to determine if some of the projections needed to be adjusted on the basis of expert judgment.

Adjustments made were based on following points:

1/ “Rules of Thumb”,
2/ Correlation among the projected variable values for a city. For example, between wages growth and retail sale growth,
3/ National macro-economic and regional forecasts.

The expected economic development for the Russian Federation as a whole for 2004 is quite positive. The forecast of the main indexes characterizing socio-economic development that were prepared by the Ministry of Economic Development and Trade of the Russian Federation for 2004, as of 19.08.03, are shown in final column of Table 5. This forecast was made before the ministry received the official annual returns for 2003 and will be re-estimated towards

greater growth, because the final results for 2003 are higher than those predicted in the middle of 2003.

In Table 5, the shaded cells contain values that the forecasting team found to be unrealistic. There were adjusted based on the points listed above to arrive at the adjusted values shown in Table 6. These values were sent to the participating cities for their review.

Case-by-case descriptions of the reasoning behind each of the adjustments are shown in notes to table 7; the table shows the changes introduced through the analysts' adjustments. First, growth rates of predicted parameters were adjusted, and after that other correspondent figures (value in constant prices and per capita) were recalculated.

Table 5 Forecast of socio-economic development for 2004
Results of prediction by model

	Kaliningrad	Tomsk	Penza	Engels	Cherepovets	Magadan	Tula	Tumen	Russia
Employment, 1000 people	242,5	188,1	243,9	88,2	140,1	47,0	306,7	214,0	66200
Employment per 1,000 population	612,5	383,4	493,2	385,7	436,0	323,7	591,3	390,4	465,9
Growth rate of employment, in % to previous year	96,5	99,8	97,0	99,2	97,3	96,1	96,3	97,3	100,5
Value of manufacturing output in price of 2001, mln. Rub.	16 864,4	16 030,5	14 863,6	6 101,1	68 218,3	4 243,9	24 393,4	11 448,0	6692900
Value of manufacturing output per capita in price of 2001, rub.	42 595,1	32 673,7	30 056,1	26 680,7	212 294,6	29 218,4	47 024,9	20 879,6	47099,6
Growth rate of manufacturing output, in % to previous year	107,4	106,4	110,7	113,5	100,7	95,2	107,3	108,1	103,6
New housing commissioned, thousands square meters	115,5	152,3	116,4	70,3	49,6	8,2	135,7	260,6	-
Average level of wages and salaries in prices of 2001, rub. per employed	6 660,1	3 057,7	3 799,0	3 150,2	6 224,3	7 786,1	4 003,5	6 241,3	4310,5
Real average level of wages and salaries in prices of 2001, in % to previous year	116,3	68,2	103,8	124,1	98,7	92,5	106,7	91,9	108,8
Value of retail sales in prices of 2001, mln. rub.	12 489,9	16 847,2	12 544,3	4 402,7	6 316,6	3 519,0	13 506,6	25 385,5	3799000
Value of retail sales per capita in prices of 2001, rub.	31 546,1	34 338,3	25 366,3	19 253,4	19 657,3	24 227,9	26 037,6	46 299,9	26734,9
Growth rate of retail sales, in % to previous year	106,3	103,8	108,8	118,5	116,1	104,8	108,3	99,9	106,9

Table 6. Forecast of socio-economic development for 2004
Results of prediction by model with adjustments – sent to cities

	Kaliningrad	Tomsk	Penza	Engels	Cherepovets	Magadan	Tula	Tumen
Employment, 1000 people	242,5	188,1	243,9	88,2	140,1	47,0	306,7	214,0
Employment per 1,000 population	612,5	383,4	493,2	385,7	436,0	323,7	591,3	390,4
Growth rate of employment, in % to previous year	96,5	99,8	97,0	99,2	97,3	96,1	96,3	97,3
Value of manufacturing output in price of 2001, mln. Rub.	16 864,4	16 030,5	14 863,6	5 645,7	68 218,3	4 323,8	24 393,4	11 448,0
Value of manufacturing output per capita in price of 2001, rub.	42 595,1	32 673,7	30 056,1	24 689,1	212 294,6	29 768,3	47 024,9	20 879,6
Growth rate of manufacturing output, in % to previous year	107,4	106,4	110,7	105,0	100,7	97,0	107,3	108,1
New housing commissioned, thousands square meters	115,5	152,3	116,4	70,3	49,6	8,2	135,7	260,6
Average level of wages and salaries in prices of 2001, rub. Per employed	6 660,1	4 358,5	3 799,0	2 665,4	6 224,3	7 786,1	4 003,5	6 723,1
Real average level of wages and salaries in prices of 2001, in % to previous year	116,3	99,0	103,8	105,0	98,7	92,5	106,7	99,0
Value of retail sales in prices of 2001, mln. Rub.	12 489,9	16 847,2	12 544,3	3 938,9	5 492,9	3 357,7	13 506,6	25 385,5
Value of retail sales per capita in prices of 2001, rub.	31 546,1	34 338,3	25 366,3	17 225,2	17 094,0	23 117,4	26 037,6	46 299,9
Growth rate of retail sales, in % to previous year	106,3	103,8	108,8	106,0	101,0	100,0	108,3	99,9

Table 7. ***Difference***

	Kaliningrad	Tomsk	Penza	Engels	Cherepovets	Magadan	Tula	Tumen
Employment, 1000 people								
Employment per 1,000 population								
Growth rate of employment, in % to previous year								
Value of manufacturing output in price of 2001, mln. rub.				-455,4		79,9		
Value of manufacturing output per capita in price of 2001, rub.				-1 991,6		549,9		
Growth rate of manufacturing output, in % to previous year				-8,5*		1,8****		
New housing commissioned, thousands square meters								
Average level of wages and salaries in prices of 2001, rub. per employed		1 300,8		-484,8				481,8
Real average level of wages and salaries in prices of 2001, in % to previous year		30,8**		-19,1*				7,1**
Value of retail sales in prices of 2001, mln. rub.				-463,8	-823,7	-161,3		
Value of retail sales per capita in prices of 2001, rub.				-2 028,2	-2 563,4	-1 110,5		
Growth rate of retail sales, in % to previous year				-12,5*	-15,1***	-4,8****		

Notes: * - The maximum growth rate of manufacturing output appears too high in Engels. The economic situation in the city has been set over the last few years, including the situation in manufacturing production. Based on several factors, lower growth is expected. Critically, the level of investment is extremely low. Therefore, the growth rate of manufacturing output was decreased 8,5 percentage points.

Parallel changes were made in other variables: decreases were made to the real average level of wages and salaries and to growth rate of retail sales, because the assigned moderate growth rate of manufacturing output could not sustain such fast growth of wages and salaries, which in turn decreases the high growth rate of retail sales.

** - The forecast of growth rate of manufacturing output in Tomsk and Tumen is 106.4% and 108.1%, respectively. On the other hand, the real average level of wages and salaries was projected at 68.2% and 91.9% of the previous year's value. Basically, growth of manufacturing output is accompanied by growth of wages and salaries. In our case we have two ways to reconcile these results: decreasing of growth rate of manufacturing output or increasing of real average level of wages and salaries. As the growth rate of manufacturing output predicted by model is quite believable in both cases, and the growth rate of wages and salaries predicted by model seems highly implausible, taking into account the previous development in these cities, macroeconomic forecast for Russia and forecast of gross regional product for both regions (Tomsk oblast and Tumen oblast), the second adjustment was chosen. The growth rate of wages and salaries was increased on 30.8 percent points (Tomsk) and 7.1 percent points (Tumen).

*** - The growth rate of retail sales in Cherepovets was decreased taking into account the forecast of average level of wages and salaries.

**** - In Magadan the growth rate of manufacturing output during last two years was quite moderate and slightly lower than average level for Russia: 2002 – 102,0% for Magadan vs. 104,3% for Russia; 2003 – 104,0% vs. 105,9%. Taking into account that the forecast for Russia for 2004 is 103,6%, the forecast for Magadan was adjusted from 95% to 97%. Decreasing of real wage was a motive for decreasing the growth rate of retail sales.

View of the Cities on the Projections Provided by IUE

This section presents two kinds of information received from the participating cities about the projections. First, some cities provided their projections of the same variables for purposes of comparison. Second, some cities offered comments on the quality and usefulness of the IUE projections. These are reviewed below on a city-by-city basis. Where a city provided its forecasts, these are presented first, beginning with the projections for Tula, shown in Table 8.

Table 8. Forecast of socio-economic indexes for 2004 in Tula

	IUE model	Forecast of Tula Administration	Deviation, %
Employment, 1000 people	306,7	330,0	-7,1
Employment per 1,000 population	591,3	634,7	-6,8
Growth rate of employment, in % to previous year	96,3	103,6	-7,3
Value of manufacturing output in price of 2001, mln. rub.	24 393,4	23 629,3	3,2
Value of manufacturing output per capita in price of 2001, rub.	47 024,9	45 449,7	3,5
Growth rate of manufacturing output, in % to previous year	107,3	104,0	3,3
New housing commissioned, thousands square meters	4 003,5	3 875,2	3,3
Real average level of wages and salaries in prices of 2001, in % to previous year	106,7	107,2	-0,5
Value of retail sales in prices of 2001, mln. rub.	13 506,6	13 064,0	3,4
Value of retail sales per capita in prices of 2001, rub.	26 037,6	24 728,4	5,3
Growth rate of retail sales, in % to previous year	108,3	106,2	2,1

Notes of Tula Administration: The deviations for most of the predicted parameters are small; levels of relative error are in an allowable interval. This shows that testing model is acceptable for prediction of the main economic parameters, characterizing socio-economic development of Tula. The difference in the employment number is explained by the fact that the administration predicted employment in economy with the use of a technique based on estimation of structural shifts in sex-age structure of population, balance of migration, and situation at labor-market in Tula.

Table 9. Forecast of socio-economic indexes for 2004 in Tomsk

	IUE model	Forecast of Tomsk Administration	Deviation, %
Employment, 1000 people	188,2	189,2	-0,5%
Employment per 1,000 population	383,6	385,8	-0,6%
Growth rate of employment, in % to previous year	99,8	100,4	-0,6%
Value of manufacturing output in price of 2001, mln. rub.	16 914,30	16 892,9	0,1%
Value of manufacturing output per capita in price of 2001, rub.	34 475,10	34 447,2	0,1%
Growth rate of manufacturing output, in % to previous year	106,0	105,9	0,1%
New housing commissioned, thousands square meters	152,3	155,0	-1,7%
Real average level of wages and salaries in prices of 2001, in % to previous year	99,1	99,9	-0,8%
Value of retail sales in prices of 2001, mln. rub.	16 156,60	16 078,0	0,5%
Value of retail sales per capita in prices of 2001, rub.	32 930,80	32 785,4	0,4%
Growth rate of retail sales, in % to previous year	104,3	106,8	-2,3%

Notes of Tomsk Administration:

The forecast prepared by IUE was attached to projections of the city budget for 2004.

Table 10. Forecast of socio-economic indexes for 2004 in Yaroslavl

	IUE model	Forecast of Yaroslavl Administration	Deviation, %
Employment, 1000 people	233,4	236,4	-1,3
Employment per 1,000 population	391,7	397,0	-1,3
Growth rate of employment, in % to previous year	98,4	101,3	-2,9
Value of manufacturing output in price of 2001, mln. rub.	44 445,2	43 368,8	2,5
Value of manufacturing output per capita in price of 2001, rub.	74 566,2	72 833,4	2,4
Growth rate of manufacturing output, in % to previous year	104,5	102,0	2,5
New housing commissioned, thousands square meters	137,3	140,0	-2,0
Real average level of wages and salaries in prices of 2001, in % to previous year	103,2	104,2	-1,0
Value of retail sales in prices of 2001, mln. rub.	14 443,9	13 726,3	5,2
Value of retail sales per capita in prices of 2001, rub.	24 232,7	23 051,9	5,1
Growth rate of retail sales, in % to previous year	110,9	105,4	5,5

Notes of Yaroslavl Administration:
No comments.

Notes of Cherepovets Administration (received by phone):
All figures are believable except growth of real average level of wages and salaries. City administration expects a more optimistic result for 2004.

Conclusions

Based on the foregoing, we draw two basic conclusions:

1. The results of projections are quite good overall, after certain expert judgments are made.
2. Most adjustments were made to projections of real wages. This suggests that this model probably should be re-estimated. It may be that the poor predictions are in part a result of including the experience of 1998 in the data set used to estimate the model.

Next Steps

These models hold some interest by themselves because they indicate that the Russian economy is operating broadly on market principles and that the data measuring economic activity are sufficiently consistent to support estimation of stochastic economic models. But the primary interest is whether use of these models will permit one to forecast future economic conditions at the municipal level with sufficient accuracy so as to be useful to local governments. In spring 2004 when data are available for 2002, we will compare forecasts made with the models for 2002 for each municipality included in the data set with the actual values using standard tests of forecasting quality.[ix] If these results are encouraging, a similar exercise will be undertaken in spring 2005 for the 2003 data. Depending on the results, we may reestimate the models, taking advantage of the longer time series.

Table 4. Results for the Complex Models

	EMPPC	MFGPC	RETAIL-PC	WAGE-LVL	NEWHSG	FIXCAP-PC
Constant	38.50 (3.22)[a]	-341.5 (-0.52)	3301.7 (6.60)	1758.6 (3.81)	-54.9 (-0.48)	5829.0 (0.74)
GDPREG-1	0.129 (2.85)		7.91 (1.74)	6.46 (5.62)		73.8 (1.73)
WAGE-LVL-1	-0.00309 (-2.27)			0.747 (17.8)		
MFGPC-1		0.993 (37.3)				
RETAIL-PC-1			0.923 (42.4)			
FIXCAP-PC-1						1.30 (5.53)
EMPPC-1	0.955 (36.2)					
CPIREG-1				-9.98 (-9.20)		
CAPBUG-1					.0072 (8.21)	-.530 (-1.97)
HSGPC-1				29.1 (1.46)	4.58 (0.68)	
NEWHSG-1					0.681 (3.78)	
NEWHSG-2					0.162 (0.76)	
INTEREST					-0.244 (-0.26)	
TEMPJLY	-1.01 (-3.18)					-397.2 (-1.08)
TEMPAVRG			122.3 (1.57)			
CENTRAL		4427.7 (4.90)		-180.9 (-1.64)		
NORTHWEST		3804.0 (4.16)		3.48 (0.03)		
SOUTH		2839.6 (3.91)		-218.4 (-1.97)		
VOLGA		3603.3 (4.16)		-176.1 (-1.69)		
URALS		2103.5 (2.54)		-357.9 (-1.96)		
SIBERIA		2468.8 (3.49)		-141.0 (-1.50)		
LRGCTY					62.0 (1.61)	-2409.7 (-1.66)
MEDCTY					5.46 (0.47)	-1626.6 (-1.45)
SMLCTY					36.7 (0.79)	8342.3 (1.63)
Dummy for 1998	-10.8 (-3.02)	-4180.9 (-8.07)	-10579.4 (-9.02)	-1747.7 (-20.0)		
N	320	320	229	320	320	320
R^2	0.903	0.962	0.904	0.918	0.674	0.659
Root MSE	19.9	3221.6	3732.3	421.1	153.4	14997
F	541.3	578.5	1073.9	200.4	752.9	47.6
Sign.	.000	.000	.000	.000	.000	.000
R^2 for model of first difference on variables other than lags[b]	.168	.330	.508	.761	.222	.156

Notes: (a) Test statistics in parentheses based on standard errors obtained using a clustered robust variance formula. (b) The dependent variable is the first difference of the variable shown at the column head and the independent variables are the same as those shown except lagged values of the dependent variable are excluded.

References

General

Berger, M.C., G.C. Blomquist, and K.Z. Sabirianova. 2002. "Compensating Differentials in Emerging Labor and Housing Markets: Estimates of Quality of Life in Russian Cities," processed.

Croushore, D. 1998. "Evaluating Inflation Forecasts." Philadelphia: Federal Reserve Bank of Philadelphia, Working Paper no. 98-14.

Diebold, F.X., andJ.A. Lopez. 1996. "Forecast Evaluation and Combination." Cambridge, MA: National Bureau of Economic Research, Technical Working Paper 192.

Granbers, A., and J. Zaitseva. 2002. "Growth Rates in the National Economic Space," *Vaprosi Economiki* [Problems of Economics], no. 9 [in Russian].

Ioffe, G., O.L. Medvedkov, Y. Medvedkov, T. Nefodova, and N. Vlasova. 2001. "Russia's Fragmented Space," in B. Ruble, J. Koehn, and N.E. Popson (eds.) *Fragmented Space in the Russian Federation.* Washington, DC: Woodrow Wilson Center Press and Johns Hopkins University Press, pp. 31-81.

Marcellino, M., J.H. Stock, and M.W. Watson. 2003. "Macroeconomic Forecasting in the Euro Area: Country Specific versus Area-wide Information," *European Economic Review*, vol. 47, pp. 1-18.

Mayo, S.K., J.I. Stein. 1995. "Housing and Labor Market Distortions in Poland: Linkages and Policy Implications," *Journal of Housing Economics*, vol. 4, pp. 153-88.

Ministry of Economics and Trade of the Russian Federation. 2002. "Regional Socio-Economic Development of the Russian Federation to 2002." http://www.economy.gov.ru/socec2002.html.

Stata Corp, *Stata Statistical Software: Release 7.0, User's Guide,* College Station, TX, 2001.

International Literature on Regional Economic Projections[x]

Bell, D.N.F. 1994. "Regional Econometric Modelling in the U.K.: A Review," *Regional Studies*, vol. 27, 1994, pp. 777-82.

Blein, U., and A. Tassinopoulos. 2001. "Forecasting Regional Employment with the ENTROP Method," *Regional Studies*, vol. 35, 2001, pp. 113-124.

Bolton, R. 1985. "Regional Econometric Models," *Journal of Regional Science*, vol. 25, no.4, pp. 495-520.

Bronars, S.G., and D.W. Jansen. 1987. "The Geographic Distribution of Unemployment Rates in the U.S.," *Journal of Econometrics*, vol. 36, pp. 251-29.

Chambers, E.J., and P.P. Talwar. 1989. "Multi-variate Time Series Forecasting Models for Alberta, British Columbia, and Manitoba: A Progress Report." Paper presented at the Ninth Annual International Symposium on Forecasting, Vancouver.

Cliff, A.D., and J.K. Ord. 1975. "Space-time Modeling with an Application to Regional Forecasting," *Transactions of the Institute of British Geographers*, vol.66, pp. 119-28.

Figlio, D.N., and B.A. Blonigen. 2000. "The Effects of Foreign Direct Investment on Local Communities," *Journal of Urban Economics*, vol. 48, pp. 338-63.

Fiske, J.R., J.C. Lamb, and M.F. Morss. 1991. "Practical Economic Forecasting for Small Regions," *Business Economics*, July, pp. 49-54.

Glickman, N. 1977. *Econometric Analysis of Regional System*. New York: Academic Press.

Miller, J.R. 1998. "Spatial Aggregation and Regional Economic Forecasting," *Annals of Regional Science*, vol. 32, pp. 253-66.

Moulton, B.R. 1990. "An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Units," *Review of Economics and Statistics*, vol.72, pp. 334-38.

Shao, G., and G.I. Treyz. 1993. "Building U.S. National and Regional Forecasting and Simulation Models," *Economic Research Systems*, vol., 5, no.1, pp. 63-77.

Talwar, P.P, and E.J. Chambers. "Forecasting Provincial Business Indicator Variables and Forecast Evaluation," *Urban Studies*, vol. 30, 1993, pp. 1763-73.

Terkla, D.G., and P.B. Doeringer. 1991. "Explaining Variations in Employment Growth, Structural and Cyclical Change among States and Local Areas," *Journal of Urban Economics*, vol. 29, pp. 329-48.

Topel, R.H. 1986. "Local Labor Markets," *Journal of Political Economy*, vol. 94, no. 3, pt.2, pp. S111-43.

Treyz, G.I. 1993. *Regional Economic Modeling: A Systematic Approach to Economic Forecasting and Policy Analysis*. Norwell, MA: Kluwer Academic.

Trivez, F.J., and J. Mur. 1999. "A Short-Term Forecasting Model for Sectoral Regional Employment," *Annals of Regional Science*, vol. 33, pp. 69-91.

[i] The major items consulted are listed in the second section of references at the end of the paper.

[ii] Berger, Blomquist, and Sabirianova (2002); Mayo (1995).

[iii] There are low simple correlations among these variables. So it is possible to include all three sets of variables in the same model.
[iv] We considered splitting the sample by size, which would ease the heteroskedasticity problem. We decided against this option, because of the sense that the data in the smaller cities maybe of relatively poor quality.
[v] An overview of this method can be found in Stata Corp (2001).
[vi] One might want to treat all of the lagged outcomes as endogenous. However, in practical terms, too few exogenous instruments are available to estimate the parameters of the resulting model.
[vii] We also experimented with use of seemingly unrelated regressions – a technique that estimates coefficients taking into account the degree of correlation of the errors across equations. The resulting estimates were relatively similar to those observed in the single equation estimates.
[viii] The short time period for which data are available also affected the estimation strategy in another way. In principle, it would be possible to estimate "h-step ahead" models where the data employed time difference between the date of the most recent data and the date for which the projection is needed. In our case there would be a three-year difference in the years for the dependent variables and the independent variables, i.e., a "3-step ahead" projection. See Marcellino, Stock and Watson (2003) for an example of applying this approach.
[ix] See, for example, Diebold and Lopez (1996) and Croushore (1998).
[x] Some references on prediction of national level CPI are also included because of the few articles available on regional CPI forecasts.

Mr. Struyk and Mr. Wissoker work at the Urban Institute in Washington, DC. Ms. Zaitsaeva works at the Institute for Urban Economics in Moscow. Work on this paper was conducted under the Cooperative Agreement, No. 118-A-00-01-00135-00, between the U.S. Agency for International Development-Russia and the Institute for Urban Economics. The authors thank German Vetrov for helpful comments during the course of the project.