NBER WORKING PAPER SERIES

LEARNING CAPITALISM THE HARD WAY--EVIDENCE FROM GERMAN REUNIFICATION

Thomas P. Triebs
Justin Tumlinson

Working Paper 19209
http://www.nber.org/papers/w19209

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 2013

We would like to thank John Morgan, Johannes Koennen, Deepak Hegde, the participants of the NBER Conference on the Economics of Religion and Culture and the participants of the Public Economics Seminar at the KU Leuven for their helpful comments. We would like to thank Steffen Elstner and Christian Seiler for assisting us with the data.The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Learning Capitalism the Hard Way--Evidence from German Reunification
Thomas P. Triebs and Justin Tumlinson
NBER Working Paper No. 19209
July 2013, Revised August 2016
JEL No. D21,D22,D84,F15,L2,N24,N64,Z1

ABSTRACT

We develop a model of firm learning in volatile markets with noisy signals and test its predictions using historical data from the Ifo Institute's Business Climate Survey. We find that firms' forecasts improve as they age. We also exploit German Reunification as a natural experiment where firms in the East are treated with ignorance about the distribution of market states. As theoretically predicted, Eastern firms make larger forecast errors than Western ones, but this gap gradually closes over the decade following Reunification. The slow convergence rate stems from differences in expectations rather than market conditions. We also find evidence for the model's predictions that improvements from learning are faster where market signals are noisier.

Thomas P. Triebs
Ifo Institute for Economic Research
at the University of Munich
Industrial Organisation and New Technologies
Poschingerstr. 5, 81679 Munich, Germany
triebs@ifo.de

Justin Tumlinson
Ifo Institute for Economic Research
at the University of Munich
Industrial Organisation and New Technologies
Poschingerstr. 5, 81679 Munich, Germany
tumlinson@ifo.de

1. Introduction

Since firms' decisions depend on their expectations of the future market state, how expectations form is central to macroeconomics and its dynamic models—market actors' aggregate decisions today determine tomorrow's economic state. For most of the last century a rich *theoretical* literature debated the relative merits of *adaptive expectations,* in which firms' backward looking predictions do not reliably lead to equilibrium prices, versus *rational expectations*, in which firms, on average, expect the true future market state. The latter model implies equilibrium but requires that market actors understand the economy better than economists themselves. To explain where such sophisticated rational expectations come from, theorists developed models in which firms learn them. Empirically though, whether firms *learn* to forecast the market or not remains open.

We show that they do. Older firms make lower forecast errors, even including firm fixed effects to rule out survivor bias as the primary explanation. But, as with individuals, firm age correlates with many (unobservable) things besides experience that could affect forecast quality—new firms are smaller, their employees tend to be younger, their markets tend to be newer and so on. To overcome this identification problem, an ideal experiment would exogenously place a cross-section of firms into a new market environment alongside otherwise similar counterparts that are very experienced in the market and compare their forecasts of subsequently shared market conditions.

German Reunification was such an event.[1] Figure 1 plots forecast error magnitudes by Western firms since 1980 and Eastern ones after Reunification. Initially, Eastern firms made much larger forecast errors than those in the West. Over time forecast errors in the East decrease and converge to Western levels. Our controlled regressions confirm the coarse implications of Figure 1 and, in so doing, empirically support a seminal, theoretical advance in macroeconomics—firms learn to forecast market conditions.

However, our empirical findings come with a practical caution—this real world convergence took a decade, despite the fact that market conditions themselves homogenized very quickly. We live in a period of geopolitical upheaval unseen since the collapse of the Iron Curtain and the Reunification of Germany. The lessons of this switch to democracy and capitalism may help set the correct expectations for the hypothetical Reunification of other countries like Korea or the political revolutions sweeping the Middle East, the accession of Turkey or departure of Britain from the EU—learning capitalism or a new market may take much longer than building its

[1] Germany was reunited on October 3rd 1990. An economic and monetary union was already established on July 1st of that year.

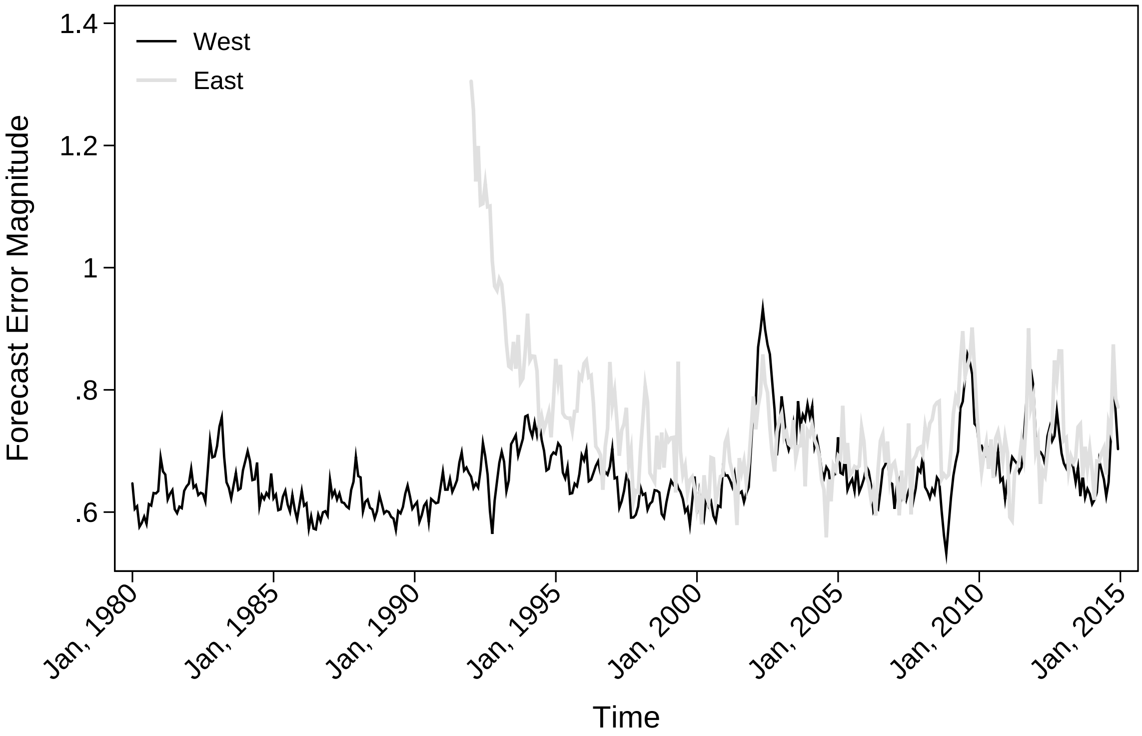


Figure 1: Forecast Errors in East and West Germany

Notes: This graph plots the monthly average forecast error magnitudes for East and West Germany. East is from 1992 on.

formal institutions.

To structure our analysis we introduce a formal model of (Bayesian) firm learning. In the model, the market state is drawn from a distribution whose parameters are initially unknown. Over time firms learn the model parameters. The model distinguishes between three types of market variance: (1) market volatility, (2) noise about signals of the future state, and (3) the imprecision in firms' subjective probability distribution of market states. We further show how these unobservable, fundamental components can be extracted from observable, agglomerated measures of uncertainty, like historical variance in firm-level states of business and forecast disagreement between industry peers. We use the model to make empirical predictions about the error magnitudes and learning as a function of these market features.

The scarcity of data across a broad cross-section of firms has generally hampered analyses of forecasting at the firm level. However, we use a unique micro data set compiled by the Ifo Institute to test our theoretical predictions. Every month since 1949 the Ifo Institute's Business Climate Survey has collected the near term expectations and assessment of the current state of business from a large cross section of German manufacturing establishments.

Our study, although one of the first to examine market forecasting and learning at the firm level empirically, is not without limitations. Since the primary purpose of the Ifo survey is to provide leading indicators of macro market health, large, established firms are oversampled, and thus our estimates of learning over firm age may not adequately capture effects in the critical infancy stage. Furthermore, although we measure the learning of Eastern firms that live through Reunification, the mechanisms remain somewhat obscure. Given that our natural experiment shocked not just Eastern firms, but the individuals and non-firm institutions of East Germany quite radically, we cannot disentangle firm level learning from societal.[2]

This paper proceeds as follows. Section 2 reviews the relevant literature. Section 3 develops our formal model. Section 4 introduces our data. Section 5 and 6 present the empirical analysis, and section 7 concludes.

2. Literature

In the macroeconomic literature, the theoretical question of how firms forecast future prices extends from at least Nerlove (1958)'s model of *adaptive expectations*, where firms simply expect next period's prices to be a weighted history of past prices, and set production quantity accordingly. Since adaptive expectations do not always converge to equilibrium, in the sense that average economic predictions may *systematically* differ from outcomes, Muth (1961) proposed the theory of *rational expectations,* under which firms' forecast errors are independently and identically distributed with zero mean in each period. In other words, no firm is systematically wrong. Although, rational expectations consistently lead to equilibria, the theory does not set out a process by which economic agents, who initially might be biased, arrive at them. Turnovsky (1969) and Cyert and DeGroot (1974) proposed a process where initially agents' subjective probability distribution differed from the objective distribution and agents updated their beliefs following Bayes' rule, eventually converging to rational expectations. Vives (1993) showed, though, that this convergence is theoretically quite slow, a significant problem for equilibrium if the underlying data generating process evolves. Related theoretical research in industrial organization showed firms converge to equilibria via learning when strategic interaction matters (see Fudenberg and Levine (1998) for a review of learning in strategic games). Our model builds upon this rich learning framework—firms' initial, (incorrect) subjective probability distributions of markets states gradually converge to the (true) objective distribution via Bayesian updating. Following the macro tradition, we consider firms to be so small that strategic interaction plays no role. Our theoretical focus is on generating empirical predictions

[2] A number of studies argue that communism imprinted individual preferences (van Hoorn and Maseland, 2010; Heineck and Süssmuth, 2010; Alesina and Fuchs-Schündeln, 2007).

about the relationship between forecast quality and several distinct types of market uncertainty, together with convergence of firm beliefs to steady state.

Empirical examinations of the *firm-level* expectation formation process are rare. In one exception, Nerlove (1983) empirically tests the expectations formation process using Ifo's Business Climate data for the period 1977/78 (an earlier subset of our data) and comparable French data. He finds that a simple error-correction model, though being "devoid of economic content" explains observed expectations "surprisingly well" (p. 1267). He proposed calculating forecast errors similar to ours, but due to computational limitations of the era, computes associations based on 2-way contingency tables for expectations and realizations instead. Bachmann and Elstner (2015) use the Ifo's Business Climate Survey data to investigate the impact of expectation biases on aggregate welfare. Despite the wide ranging theoretical literature on *learning* in expectation formation, empirical evidence is limited to a recent working paper by Doraszelski et al. (2014).[3] They show that after the liberalization of the previously regulated market for (electric power) frequency response in the UK, competing electricity suppliers' prices converged to the Nash equilibrium over three years. Ours is, as far as we know, the first multi-industry examination of firm-level forecasting using controlled regressions, and thus the first to disentangle the role of market features in learning.

3. Theoretical Models

To structure our thinking we present stylized models of our data generating process under (1) pure rational expectations and (2) learning. The two models are identical, except that in the learning model, firms do not know one of the distributional parameters of the data generating process—they learn it over time. These simple models, like the classical models of learning in rational expectations developed by Townsend (1978) and Feldman (1987) and extended by many others, (unrealistically) treat states as serially independent, and assume (improbably) that firms know the distributional family of the underlying data generating processes. More generally, these models should not be taken as trying to capture the complexities of the true forecasting procedure that firms execute in predicting their future business conditions, but rather to parsimoniously focus on the intuitive difference between pure rational expectations and learning models under markets exhibiting various types of uncertainty.

At the beginning of period t, nature draws two hidden states relevant for firm i's change in

[3] A literature on professional forecasters examines biases due to career concerns and sponsoring institutions (e.g. Cho and Hersch, 1998), as well as individual-level correlates of forecasting ability (e.g. Kim et al., 2011) but not learning.

business conditions S_{it} (literally *Geschäftslage*): X_{it} is an idiosyncratic firm state distributed $Normal\left(\mu_{X_i}, \sigma^2_{X_i}\right)$, and Y_t is a market state common to all firms distributed $Normal\left(\mu_Y, \sigma^2_Y\right)$, where X_{it} and Y_t are independent. For simplicity, assume firm level business conditions are simply the sum of these variables: $S_{it} = X_{it} + Y_t$. The information available at the beginning of period t includes all *previous* state realizations $\Omega_{it} = \{X_{it-1}, \ldots, X_{i1}, Y_{t-1}, \ldots, Y_1\}$ and a firm specific signal of the current market state $\hat{Y}_{it} = Y_t + \varepsilon_{it}$, where noise ε_{it} is, for all firms, distributed $Normal\left(0, \sigma^2_\varepsilon\right)$. Thus, $X_{it} | \hat{Y}_{it}, \Omega_{it} = X_{it}$ and $Y_t | \hat{Y}_{it}, \Omega_{it}$ are still independent. The key difference between the *rational expectations* model and the *learning model* is that in the latter, firms do not know the value of μ_Y; they must learn it. Under the learning model, firm i holds prior beliefs about the mean of the market state variable, but these beliefs are updated over time. We assume the prior beliefs of firm i about μ_Y are normally distributed with mean μ_{iY0} and variance σ^2_{iY0}.

Then, the firm makes a prediction about its state of business equal to the sum of conditional forecasts about its idiosyncratic and market states.

$$\bar{S}_{it} = \bar{X}_{it} + \bar{Y}_{it} = E\left[X_{it}\right] + E\left[Y_t | \hat{Y}_{it}, \Omega_{it}\right]$$

At the end of period t the realized state variables are revealed, and a directional forecast error $S_{it} - \bar{S}_{it}$ is computed. A positive value indicates that the firm was pessimistic—it predicted a worse change in business state than actually occurred. A negative value indicates that the firm was optimistic—it predicted a better change in state of business than actually occurred. We are interested in the expected *magnitude* of this error or so-called *mean squared error*

$$\begin{aligned}
E\left[\left(S_{it} - \bar{S}_{it}\right)^2\right] &= E\left[\left(\left(X_{it} + Y_t\right) - \left(E\left[X_{it}\right] + E\left[Y_t | \hat{Y}_{it}, \Omega_{it}\right]\right)\right)^2\right] \\
&= E\left[\left(X_{it} - E\left[X_{it}\right]\right)^2\right] + E\left[\left(Y_t - E\left[Y_t | \hat{Y}_{it}, \Omega_{it}\right]\right)^2\right] \\
&+ 2E\left[\left(X_{it} - E\left[X_{it}\right]\right)\left(Y_t - E\left[Y_t | \hat{Y}_{it}, \Omega_{it}\right]\right)\right] \\
&= Var\left[X_{it}\right] + Var\left[Y_t | \hat{Y}_{it}, \Omega_{it}\right]
\end{aligned}$$

where the independence of X_{it} and Y_t implies

$$E\left[\left(X_{it} - E\left[X_{it}\right]\right)\left(Y_t - E\left[Y_t | \hat{Y}_{it}, \Omega_{it}\right]\right)\right] = Cov\left(X_{it}, Y_t | \hat{Y}_{it}, \Omega_{it}\right) = 0$$

The following analysis computes comparative statics on the error magnitude with respect to time, the variance of the market state and signal noise variables. Then these statics are translated to empirical predictions on how error magnitude changes with respect to empirically observable quantities: time, the computed market volatility, and the computed signal noise across industries.

3.1. Rational Expectations

When the distribution parameters of Y_t are known, history Ω_{it} is irrelevant to making predictions about the future, but current market signals $\hat{Y}_{it}$ are quite useful. It is well-known that the mean and variance of correct (normally distributed) posterior beliefs over Y_t are

$$\bar{Y}_{it} = E\left[Y_t | \hat{Y}_{it}, \Omega_{it}\right] = \frac{\sigma_\varepsilon^2}{\sigma_Y^2 + \sigma_\varepsilon^2}\mu_Y + \frac{\sigma_Y^2}{\sigma_Y^2 + \sigma_\varepsilon^2}\hat{Y}_{it} \tag{1}$$

$$\bar{\sigma}_{iYt}^2 = Var\left[Y_t | \hat{Y}_{it}, \Omega_{it}\right] = \frac{\sigma_Y^2\sigma_\varepsilon^2}{\sigma_Y^2 + \sigma_\varepsilon^2} \tag{2}$$

respectively. Thus,

$$E\left[\left(S_{it} - \bar{S}_{it}\right)^2\right] = Var\left[X_{it}\right] + Var\left[Y_t | \hat{Y}_{it}, \Omega_{it}\right] = \sigma_X^2 + \frac{\sigma_Y^2\sigma_\varepsilon^2}{\sigma_Y^2 + \sigma_\varepsilon^2} \tag{3}$$

In this case, we can generate the following comparative statics on MSE under rational expectations: (1) the expected magnitude of forecast errors does not change over time,

$$\frac{d}{dt}E\left[\left(S_{it} - \bar{S}_{it}\right)^2\right] = 0$$

(2) the expected magnitude of forecast errors increases with the noisiness of market signals,

$$\frac{d}{d\sigma_\varepsilon^2}E\left[\left(S_{it} - \bar{S}_{it}\right)^2\right] = \left(\frac{\sigma_Y^2}{\sigma_Y^2 + \sigma_\varepsilon^2}\right)^2 > 0$$

and (3) the expected magnitude of forecast errors increases with the volatility of the market,

$$\frac{d}{d\sigma_Y^2}E\left[\left(S_{it} - \bar{S}_{it}\right)^2\right] = \left(\frac{\sigma_\varepsilon^2}{\sigma_Y^2 + \sigma_\varepsilon^2}\right)^2 > 0$$

3.2. Learning

It is well known that the posterior predictive distribution of Y_t, normally distributed with unknown mean and variance σ_Y^2, *unconditional on the signal* is $Normal\left(\mu_{iYt}, \sigma_{iYt}^2 + \sigma_Y^2\right)$ where

$$\mu_{iYt} = \frac{\sigma_Y^2}{t\sigma_{iY0}^2 + \sigma_Y^2}\mu_{i0} + \frac{t\sigma_{iY0}^2}{t\sigma_{iY0}^2 + \sigma^2}\bar{Y}$$

$$\sigma_{iYt}^2 = \frac{\sigma_{iY0}^2\sigma_Y^2}{t\sigma_{iY0}^2 + \sigma_Y^2}$$

and $\bar{Y}$ is the sample mean of realized market states up to time t.[4, 5] Thus, substituting μ_{iYt} and $\sigma^2_{iYt} + \sigma^2_Y$ for μ_Y and σ^2_Y in equations (1) and (2) of the previous subsection, the mean and variance of the normally distributed posterior predictives over Y_t are

$$\bar{Y}_{it} = E\left[Y_t | \hat{Y}_{it}, \Omega_{it}\right] = \frac{\sigma^2_\varepsilon}{(\sigma^2_{iYt} + \sigma^2_Y) + \sigma^2_\varepsilon}\mu_{iYt} + \frac{(\sigma^2_{iYt} + \sigma^2_Y)}{(\sigma^2_{iYt} + \sigma^2_Y) + \sigma^2_\varepsilon}\hat{Y}_{it}$$
$$\bar{\sigma}^2_{iYt} = Var\left[Y_t | \hat{Y}_{it}, \Omega_{it}\right] = \frac{(\sigma^2_{iYt} + \sigma^2_Y)\,\sigma^2_\varepsilon}{(\sigma^2_{iYt} + \sigma^2_Y) + \sigma^2_\varepsilon}$$

respectively. Thus, we can generate the following comparative statics on the MSE under learning by substituting $\sigma^2_{iYt} + \sigma^2_Y$ for σ^2_Y into eqn. (3) and taking derivatives. (1) The expected magnitude of forecast errors decreases over time:

$$\frac{d}{dt}E\left[\left(S_{it} - \bar{S}_{it}\right)^2\right] = -\sigma^2_Y\left(\frac{\sigma^2_{iY0}\sigma^2_\varepsilon}{t\sigma^2_{iY0}\left(\sigma^2_Y + \sigma^2_\varepsilon\right) + \sigma^2_Y\left(\sigma^2_{iY0} + \sigma^2_Y + \sigma^2_\varepsilon\right)}\right)^2 < 0 \tag{4}$$

(2) The expected magnitude of forecast errors increases with the noisiness of market signals:

$$\frac{d}{d\sigma^2_\varepsilon}E\left[\left(S_{it} - \bar{S}_{it}\right)^2\right] = \left(\frac{\left(t\sigma^2_{iY0} + \sigma^2_{iY0} + \sigma^2_Y\right)\sigma^2_Y}{t\sigma^2_{iY0}\left(\sigma^2_Y + \sigma^2_\varepsilon\right) + \sigma^2_Y\left(\sigma^2_{iY0} + \sigma^2_Y + \sigma^2_\varepsilon\right)}\right)^2 > 0 \tag{5}$$

(3) The expected magnitude of forecast errors increases with the volatility of the market:

$$\frac{d}{d\sigma^2_Y}E\left[\left(S_{it} - \bar{S}_{it}\right)^2\right] = \frac{\left(\left(t\sigma^2_{iY0} + \sigma^2_Y\right)\sigma^2_\varepsilon\right)^2 + t\left(\sigma^2_{iY0}\sigma^2_\varepsilon\right)^2}{\left(t\sigma^2_{iY0}\left(\sigma^2_Y + \sigma^2_\varepsilon\right) + \sigma^2_Y\left(\sigma^2_{iY0} + \sigma^2_Y + \sigma^2_\varepsilon\right)\right)^2} > 0 \tag{6}$$

Observe that these comparative statics converge precisely to those under rational expectations as $t \to \infty$. That is to say, our learning model, like many others before ours, converges to rational expectations with experience.

By taking the derivative of these statics with respect to time, we can also compute how learning evolves with experience. (1) Learning slows over time, because the influence of each new piece of information is smaller relative to the accumulated stock of knowledge:

$$\frac{d^2}{dt^2}E\left[\left(S_{it} - \bar{S}_{it}\right)^2\right] = \frac{2\sigma^2_{iY0}\sigma^2_Y\left(\sigma^2_Y + \sigma^2_\varepsilon\right)\left(\sigma^2_{iY0}\sigma^2_\varepsilon\right)^2}{\left(t\sigma^2_{iY0}\left(\sigma^2_Y + \sigma^2_\varepsilon\right) + \sigma^2_Y\left(\sigma^2_{iY0} + \sigma^2_Y + \sigma^2_\varepsilon\right)\right)^3} > 0 \tag{7}$$

[4] The posterior predictive distribution is the distribution of unobserved observations, conditional on the observed data.

[5] See https://www.cs.ubc.ca/~murphyk/Papers/bayesGauss.pdf for a derivation.

(2) The magnitude of forecast errors diminishes more quickly in markets with noisier signals

$$\frac{d^2}{dtd\sigma_\varepsilon^2}E\left[\left(S_{it}-\bar{S}_{it}\right)^2\right]=-\frac{2\left(t\sigma_{iY0}^2+\sigma_{iY0}^2+\sigma_Y^2\right)\sigma_\varepsilon^2\left(\sigma_{iY0}^2\sigma_Y^2\right)^2}{\left(t\sigma_{iY0}^2\left(\sigma_Y^2+\sigma_\varepsilon^2\right)+\sigma_Y^2\left(\sigma_{iY0}^2+\sigma_Y^2+\sigma_\varepsilon^2\right)\right)^3}<0 \tag{8}$$

The intuition for $\frac{d^2}{dtd\sigma_\varepsilon^2}E\left[\left(S_{it}-\bar{S}_{it}\right)^2\right]<0$ is somewhat subtle. Early on, *i.e.* while a firm is very uncertain about the learned parameter, it must rely heavily on the signal of the current market state to forecast. Thus, early on, firms in industries with poor quality signals do very poorly relative to those in industries with better signals. But, as time progresses, firms learn the mean of the market state distribution, and thus, become less reliant on the signals alone. Hence this learning is more important to firms in industries with noisy signals.

(3) Although readily computed, the qualitative impact of market volatility is ambiguous:

$$\frac{d^2}{dtd\sigma_Y^2}E\left[\left(S_{it}-\bar{S}_{it}\right)^2\right]=\frac{\left[t\sigma_{iY0}^2\left(\sigma_Y^2-\sigma_\varepsilon^2\right)+\sigma_Y^2\left(\sigma_{iY0}^2+3\sigma_Y^2+\sigma_\varepsilon^2\right)\right]\left(\sigma_{iY0}^2\sigma_\varepsilon^2\right)^2}{\left(t\sigma_{iY0}^2\left(\sigma_Y^2+\sigma_\varepsilon^2\right)+\sigma_Y^2\left(\sigma_{iY0}^2+\sigma_Y^2+\sigma_\varepsilon^2\right)\right)^3} \tag{9}$$

The sign of (9) turns on the square bracketed factor in the numerator, which depend on the relative magnitude of market volatility σ_Y^2 and signal noise σ_ε^2. However, in industries, which are very volatile relative to the noisiness of their signals $\left(i.e.\ \sigma_Y^2>\sigma_\varepsilon^2\right)$, it is intuitive that learning will help less as the comparative static suggests. In such industries, firms must always rely heavily on the market signal, even after the learned parameter is known with certainty. Thus, learning is of limited value in such industries. This intuition is reflected in the complete derivative.

3.3. Mapping to Data

Although we know of no other study using microdata to test whether and how firms learn to forecast under uncertainty, several types of sample variance found in firm level forecast data have been used to quantify the macro uncertainty facing markets or the economy generally. In particular, others have used (1) longitudinal variation in realized states of business (e.g. Comin and Mulani, 2006), and (2) contemporaneous variation in individual firm predictions about the future state of business to characterize market uncertainty (Bachmann et al., 2013; Bloom, 2014). We will also take advantage of this variation to test the predictions of our learning model; however, neither of these sample variances generically capture the independent variables of our comparative statics: σ_Y^2 or σ_ε^2.

In our data, we observe every firm's realized state of business S_{it} and associated forecast $\bar{S}_{it}$ in every period. We do not, though, directly observe market state Y_t or each firm's signal of it $\hat{Y}_{it}$. Thus, we cannot directly observe σ_Y^2 or σ_ε^2. However, we can observe quantities, which under some additional structure asymptote to these unobservable parameters. In particular, we assume industries are large enough that the *average* idiosyncratic firm state does not vary from period to period. This is reasonable since, by definition, individual firm states are idiosyncratic and independent from one another; with many firms, individual deviations from their individual means wash out.

First, we argue that under this assumption, the longitudinal sample variance of the realized states of business vol_t is a sufficient statistic for volatility of the unobserved (by the econometrician) state of the market σ_Y^2. We measure vol_t as follows:

$$\begin{aligned} vol_t &= \frac{1}{T}\sum_{x=t}^{T}\left(\frac{1}{n}\sum_{i=1}^{n}S_{ix} - \frac{1}{T}\sum_{\tau=t}^{T}\left(\frac{1}{n}\sum_{j=1}^{n}S_{j\tau}\right)\right)^2 \\ &= \frac{1}{T}\sum_{x=t}^{T}\left(\frac{1}{n}\sum_{i=1}^{n}S_{ix} - \frac{1}{T}\sum_{\tau=t}^{T}\left(\frac{1}{n}\sum_{j=1}^{n}(X_{j\tau}+Y_\tau)\right)\right)^2 \\ &= \frac{1}{T}\sum_{x=t}^{T}\left(E\left[X_{ix}\right] + Y_x - \frac{1}{T}\sum_{\tau=t}^{T}\left(E\left[X_{i\tau}\right]+Y_\tau\right)\right)^2 \end{aligned}$$

Assuming n is large, by the Strong Law of Large Numbers (SLLN), $E\left[X_{it}\right] \to \mu_X$. Hence,

$$vol_t = \frac{1}{T}\sum_{x=t}^{T}\left(Y_x - \frac{1}{T}\sum_{\tau=t}^{T}Y_\tau\right)^2 = \sigma_Y^2$$

Notice that the second form of variation used in the literature, namely disagreement in predictions, is really an agglomeration of at least three distinct sources of uncertainty: (1) market state volatility σ_Y^2, (2) signal noise σ_ε^2, and (3) imprecision of beliefs about the learned parameter σ_{iYt}^2. For example, two firms may receive identical signals in terms of both content and noise, but they may incorporate them very differently in their forecast depending on the relative volatility of their respective markets (or the confidence they have in their market understanding)—a firm in a volatile market (or an inexperienced one) will weight its forecast toward the information in the signal since its priors are less informative. Thus, although disagreement, dis_t, may indeed be a generic measure of uncertainty, suitable in some empirical settings, it is too coarse for the more precise forecasting model and associated comparative statics we have described above.

We can extract σ_ε^2 for each industry from our observed measure of disagreement dis_t, using vol_t and some assumptions. We measure disagreement, or the variance in firms' predictions of their states of business as follows

$$
\begin{aligned}
dis_t &= \frac{1}{n}\sum_{i=1}^{n}\left(\bar{S}_{it}-\frac{1}{n}\sum_{j=1}^{n}\bar{S}_{jt}\right)^2\\
&= \frac{1}{n}\sum_{i=1}^{n}\left(\left(E\left[X_{it}\right]+E_i\left[Y_t|\,\hat{Y}_{it},\Omega_{it}\right]\right)-\frac{1}{n}\sum_{j=1}^{n}\left(E\left[X_{jt}\right]+E\left[Y_t|\,\hat{Y}_{jt},\Omega_{jt}\right]\right)\right)^2
\end{aligned}
$$

Assuming n is large, by the SLLN $E\left[X_{it}\right]\to\mu_X$, and thus

$$
dis_t = \frac{1}{n}\sum_{i=1}^{n}\left(E\left[Y_t|\,\hat{Y}_{it},\Omega_{it}\right]-\frac{1}{n}\sum_{j=1}^{n}E\left[Y_t|\,\hat{Y}_{jt},\Omega_{jt}\right]\right)^2
$$

where firm i 's expectation of the future market state is given by

$$
E\left[Y_t|\,\hat{Y}_{it},\Omega_{it}\right] = \frac{\sigma_\varepsilon^2}{\left(\sigma_{iYt}^2+\sigma_Y^2\right)+\sigma_\varepsilon^2}\mu_{iYt}+\frac{\left(\sigma_{iYt}^2+\sigma_Y^2\right)}{\left(\sigma_{iYt}^2+\sigma_Y^2\right)+\sigma_\varepsilon^2}\left(Y_t+\varepsilon_{it}\right)
$$

Notice that the relative weight a firm puts on its priors about the firm state relative to the weight the firm puts on its signal depends not just on the volatility of the market state σ_Y^2 relative to the noisiness of the signal σ_ε^2, but also on how sure the firm is that its estimate of the learned parameter is correct σ_{iYt}^2. Since the imprecision of beliefs should only influence the predictions of Eastern firms we measure industry level disagreement (and signal noise) using Western firms only, who are confident in their understanding of the model. Hence, we take $\sigma_{iYt}^2\to 0$:

$$
E\left[Y_t|\,\hat{Y}_{it},\Omega_{it}\right] = \frac{\sigma_\varepsilon^2}{\sigma_Y^2+\sigma_\varepsilon^2}\mu_{Yt}+\frac{\sigma_Y^2}{\sigma_Y^2+\sigma_\varepsilon^2}\left(Y_t+\varepsilon_{it}\right)
$$

so that disagreement may be simply written as a function of volatility and signal noise:

$$
\begin{aligned}
dis_t &= \frac{1}{n}\sum_{i=1}^{n}\left(\frac{\sigma_\varepsilon^2}{\sigma_Y^2+\sigma_\varepsilon^2}\mu_{Yt}+\frac{\sigma_Y^2}{\sigma_Y^2+\sigma_\varepsilon^2}\left(Y_t+\varepsilon_{it}\right)-\frac{1}{n}\sum_{j=1}^{n}\left(\frac{\sigma_\varepsilon^2}{\sigma_Y^2+\sigma_\varepsilon^2}\mu_{Yt}+\frac{\sigma_Y^2}{\sigma_Y^2+\sigma_\varepsilon^2}\left(Y_t+\varepsilon_{jt}\right)\right)\right)^2\\
&= \frac{1}{n}\sum_{i=1}^{n}\left(\frac{\sigma_Y^2}{\sigma_Y^2+\sigma_\varepsilon^2}\varepsilon_{it}-\frac{1}{n}\sum_{j=1}^{n}\left(\frac{\sigma_Y^2}{\sigma_Y^2+\sigma_\varepsilon^2}\varepsilon_{jt}\right)\right)^2\\
&= \left(\frac{\sigma_Y^2}{\sigma_Y^2+\sigma_\varepsilon^2}\right)^2\frac{1}{n}\sum_{i=1}^{n}\left(\varepsilon_{it}-\frac{1}{n}\sum_{i=1}^{n}\varepsilon_{it}\right)^2\\
&= \left(\frac{\sigma_Y^2}{\sigma_Y^2+\sigma_\varepsilon^2}\right)^2\sigma_\varepsilon^2
\end{aligned}
$$

Thus, since under our assumptions, we know σ_Y^2 , it is straightforward to solve for σ_ε^2

$$dis_t\left(\left(\sigma_Y^2\right)^2 + 2\sigma_Y^2\sigma_\varepsilon^2 + \left(\sigma_\varepsilon^2\right)^2\right) = \left(\sigma_Y^2\right)^2\sigma_\varepsilon^2$$
$$dis_t\left(\sigma_\varepsilon^2\right)^2 + \sigma_Y^2\left(2dis_t - \sigma_Y^2\right)\sigma_\varepsilon^2 + dis_t\left(\sigma_Y^2\right)^2 = 0$$

or using the quadratic formula

$$\sigma_\varepsilon^2 = \frac{-\sigma_Y^2\left(2dis_t - \sigma_Y^2\right) + \sqrt{\left(2dis_t - \sigma_Y^2\right)^2 - \left(2dis_t\sigma_Y^2\right)^2}}{2dis_t}$$

Thus, under our assumptions we can compute proxies for unobservable market volatility and market signal noise from the volatility of reported firm business conditions and the disagreement in their predictions about future business conditions. These computed measures are used as independent variables in our empirical tests.

4. Data

4.1. Forecast Errors

No doubt, the paucity of panel data about what firms think of future market conditions has hampered empirical studies of firm-level forecasting. We test our predictions using data from the Ifo Institute's Business Climate Survey, which, to our best knowledge, is the oldest survey on firm level, market expectations and realizations in existence. Since the data was collected to provide leading indicators for the German economy rather than academic research, its use presents some addressable challenges. The Ifo Institute began surveying firms in the Federal Republic of Germany in November 1949; firms from former East Germany were added beginning in 1991. Our sample is monthly from 1980 to 2014 for West Germany and from 1992 to 2014 for former East Germany. We drop 1991 observations for the East, because administrative difficulties render these earliest Eastern observations unreliable. The data is collected at the product level. Initially, in 1980, we have more than 4000 products in the cross-section. By the end, in 2014, about 2500 products remain in our sample.[6] Our sample includes only manufacturing firms. Following Nerlove (1983, footnote 15) we treat product-level observations as independent; there are very few multi-product firms. The panel is unbalanced: firms enter, exit, and occasionally do not respond to the survey. Table 1 provides summary statistics for our variables by region.

Like most surveys of this kind, the responses are granular, in our case trichotomous (*i.e.* '+',

[6]For comparison, the Business Outlook Survey of the Federal Reserve Bank of Philadelphia surveys about 100 to 125 large manufacturing firms every month.

'=' or '-'). But since the response categories are ordered, we can compare forecasts and realizations to construct forecast errors. However, a mismatch between time horizons for expectations and realizations, as well as the fact that expectations are cast as changes but realizations are in levels complicates the forecast error calculation. We now describe our calculation of forecast errors over the state of business (*Geschäftslage*).

The survey question regarding the realization of the current state of business reads:

> *Current situation: We assess our business state for product X to be (a) good, (b) satisfactory, or (c) bad.*

And the question about the expectation reads:

> *Expectations for the next six months: Our business state for product X will be (a) better, (b) about the same, or (c) worse.*

Taking the answer to the former question as the current business state $S_t \in \{l = -1, m = 0, h = +1\}$, the answer to the latter question as the prediction of future state of business, expressed as a change from the current state, $\Delta_t \in \{``-", ``=", ``+"\}$, and the average states of business over the next six months as the realization $R_t = \frac{1}{6}\sum_{i=t+1}^{t+6} S_i$ we construct a quantitative forecast error measure using the following parameterizable formula:

$$\varepsilon_t\left(S, \Delta, R\right) = \begin{cases} +2 & \text{if} \quad (\Delta = ``-") \wedge \left(R \in \left[H^S, 1\right]\right) \\ +1 & \text{if} \quad \left[\begin{array}{c} \left((\Delta = ``-") \wedge \left(R \in \left(L^S, H^S\right)\right)\right) \\ \vee \left((\Delta = ``=") \wedge \left(R \in \left[H^S, 1\right]\right)\right) \end{array}\right] \\ 0 & \text{if} \quad \left[\begin{array}{c} \left((\Delta = ``-") \wedge \left(R \in \left[-1, L^S\right]\right)\right) \\ \vee \left((\Delta = ``=") \wedge \left(R \in \left(L^S, H^S\right)\right)\right) \\ \vee \left((\Delta = ``+") \wedge \left(R \in \left[H^S, 1\right]\right)\right) \end{array}\right] \\ -1 & \text{if} \quad \left[\begin{array}{c} \left((\Delta = ``=") \wedge \left(R \in \left[-1, L^S\right]\right)\right) \\ \vee \left((\Delta = ``+") \wedge \left(R \in \left(L^S, H^S\right)\right)\right) \end{array}\right] \\ -2 & \text{if} \quad (\Delta = ``+") \wedge \left(R \in \left[-1, L^S\right]\right) \end{cases}$$

We assume the following parameter restrictions: for all S, $-1 \leq L^S \leq H^S \leq 1$, $H^m = -L^m$ and $H^h = -L^l$. Thus, a complete parameterization is defined by H^h, H^m and H^l.

The following intuition lies behind the formula. Relative to each current state of business, the space of average future realizations is divided into three sequential intervals, which correspond to the prediction possibilities: *better*, the *same* and *worse*. The span of these intervals varies depending on the current state. For example, if the current state is h then reaching a *better* state requires surpassing a much higher threshold than if the current state is l. Then the formula computes 0 error if the predicted change Δ matches the interval into which average future realizations fall, a $+1$ error if the predicted change matches the interval one below the one in which average future realizations fall, a $+2$ error if the predicted change matches the interval two below the one in which average future realizations fall, and analogously for negative errors. This means that a firm in a h state, predicting neutral ("=") change, will make a $+1$ error if average future realizations falls in the interval $\left[H^h, 1\right]$, say if all future realizations were h. One could argue that this is not an error since we do not really know if observing the next six future states in a row as h really means that the firm's state of business improved—it was, after all, h to begin with. To rule this out as an error one would set $H^h = 1$. On the other hand, reporting the next six states of business as h is unusual even for a firm currently reporting an h state of business and could reasonably be interpreted as an improvement in business state. By setting H^h and H^l to more moderate levels, one could, in theory, capture some of this information. In our main specifications we use the following parameterization: $H^h = \frac{2}{3}$, $H^m = \frac{1}{3}$ and $H^l = 0$. However, in the appendix we show that our results are robust to a number of parameter choices.

Could low forecast errors be an artifact of firms reporting neutral expectations and neutral realizations, simply as a lazy default for the survey respondent? About 57 percent of firm expectations are neutral. Reporting a neutral expectation is not completely random. Estimates from a probability model show that older firms, larger firms, and firms that have responded frequently before are more likely to give a neutral expectation. Firms that self-report being constrained and firms in markets with higher import penetration are less likely to give a neutral response, as are those in more volatile markets and firms facing higher signal noise. There is no statistically significant difference between East and West or exporters and non-exporters with respect to forecasting neutral changes. Table 5 in the appendix gives the regression results. We further show that our results are robust to excluding observations in which firms report neutral expectations.

4.2. Market Attributes

We characterize the market by its current state, its volatility, its signal noise and import penetration. First, we proxy for the current market state as the industry-period average response on the survey question about the current state of business (*Geschäftslage*) as described in Section 3.3. Although individual firms report the state of business as good, satisfactory, or bad our computed market state variable is continuous. We define industries at the two-digit level (using the German WZ 2008 classification).

Volatility describes that variance in market states. Ideally, we would measure current volatility (σ_Y^2) over the distribution of all possible states that the market can assume. But, of course, in a cross-section only the realized market state is observed and not the distribution. Therefore, we proxy for volatility as the time series variance of the first differences of the average state of business for rolling, forward looking, 4 year windows.[7] Making the window forward-looking permits a volatility measure (vol) in the East immediately after Reunification, but it means we have to truncate our sample at 2013, which is of little consequence, since the regions have converged by then.

Market signal noise (σ_ε^2) is calculated from market volatility (σ_Y^2) and disagreement (dis) using our theoretical structure as described in Section 3.3. Disagreement is the industry-level variance over firm-level forecasts for the state of business.

Import penetration is given by the logarithm of the ratio of imports over production at the industry level. The data is available in OECD's STAN database of industrial production from 1990 only. As shown in Table 1 import penetration is lower in the West.

[7] The windows are necessarily shorter for the last four years of the sample. The results do not qualitatively change when dropping the four final years completely.

Table 1: Summary Statistics

	West				East			
	Mean	*SD*	*Max.*	*Min.*	*Mean*	*SD*	*Max.*	*Min.*
Squared Forecast Error	0.6532	0.920	4.00	0.00	0.7536	1.033	4.00	0.00
Market Volatility	0.0038	0.004	0.09	0.00	0.0049	0.010	0.28	0.00
Signal Noise	0.9900	0.011	1.00	0.57	0.9890	0.011	1.00	0.77
Market State	−0.0420	0.256	0.86	−1.00	−0.0612	0.256	1.00	−1.00
Firm constrained (=1)	0.3511	0.477	1.00	0.00	0.4740	0.499	1.00	0.00
# Employees	484.6706	2777.185	120000.00	0.00	171.2324	443.815	11825.00	0.00
Firm age	64.1477	51.241	403.00	0.00	24.5048	38.156	303.00	0.00
Exporter (=1)	0.7765	0.417	1.00	0.00	0.5821	0.493	1.00	0.00
Import penetration	−1.4489	0.799	0.93	−3.06	−1.3857	0.783	0.93	−3.06
Observations	1314038				157710			

4.3. Firm attributes

Eastern industries were broadly restructured when the market economy was introduced. The communist East had neither private property nor capitalist markets. Production units were part of large *Kombinate* that replaced market transactions by bureaucracy. As we see in Table 1 firms in the East are smaller, despite having been much more integrated before Reunification. In the West the average firm has about 480 employees but in the East only about 170 (though as Figure 2 shows this size difference stems mostly from a few very large Western firms). For the whole sample, the size distribution is skewed; about 75 percent of firms have less than 290 production employees. Eastern firms, on average, are only half as old as Western firms (possibly due to a measurement problem for age as discussed below). East German firms, though relatively well endowed compared to firms in other communist countries, had outdated capital equipment and despite a high level of formal education, employee skills did not suit a modern market economy and its division of labor (Fritsch and Mallok, 1998).

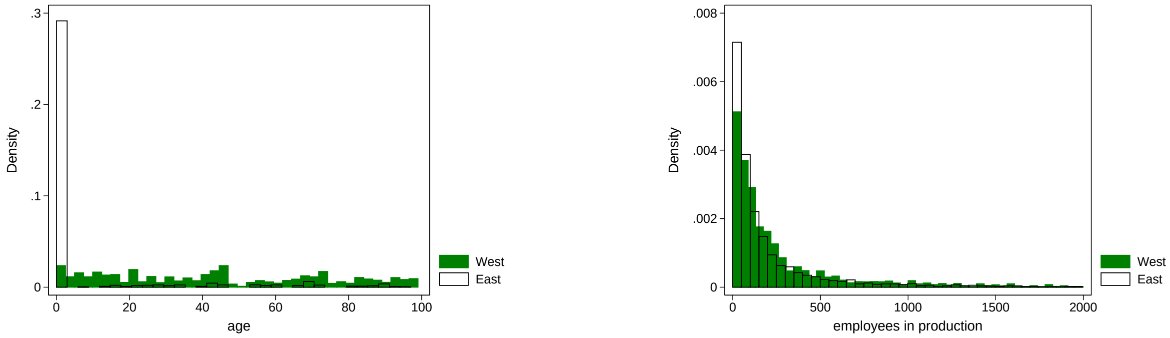
Firms report whether they are (capital or labor) "constrained". On average 47 percent of firms in the East report being constrained, but only about 35 percent of Western firms do. Since this question is asked only quarterly, except for the East in the years 1991 to 2000 where it appeared monthly, we assume the constraint or lack thereof lasts the entire quarter. Previous studies report that at Reunification the physical productivity of Eastern plants was at most 50 percent of the productivity of comparable Western manufacturing plants (Fritsch and Mallok, 1998). However, by the time our Eastern sample starts in 1992 the lowest productivity plants had already exited and most other plants had been privatized. About 77 percent of Western firms export, while only 58 percent of Eastern firms do.

Level differences in the productivity of firms, whether between East and West or simply across the street, should not obviously lead to differences in forecasts, since these idiosyncratic firm capabilities are well-known within each firm making their own forecasts.

4.3.1. Firm Age

In our theory firms learn with experience. A natural proxy for experience is age, but our age data has several shortcomings. In the West, the median firm age is 56 years and the median age of first time respondents is 41 years, older than one would expect to find in a random sample of firms. There are at least two plausible explanations for this deviation. First, because we collected the foundation date more recently than the business state/forecast information, it is likely biased by survivorship.[8] Second, given the survey's purpose, it may over sample mature

[8] We obtain the age variable from a recent search of the Amadeus database and a recent Ifo survey.



(a) Age **(b)** Size

Figure 2: Eastern vs. Western Firm Age and Size

Notes: The left panel plots the densities of age for East and West in 1992. The right panel plots the densities of firm size (number of employees) for East and West in 1992. We truncate both distributions at 2000 production employees; less than 3% of observations have more.

firms.

In the East, our firm age data has an additional problem. A disproportionate number report foundation coincident with Reunification, which is likely an artifact of privatization and does not reflect the maturity of the firm, its employees or business. This view is supported by the fact that although the age distributions between Eastern firms and Western firms differ significantly immediately following Reunification in 1992, there is little difference in the firm size distributions between the two regions (see Figure 2). In the East there is a spike in the number of new firms in 1992.

5. Empirical Analysis

Our theoretical model assumes that (some) firms do not completely understand the process, which generates future states of business, but it is agnostic about why—learning proceeds in the same way, regardless of the cause of the imperfection in understanding. In this section, we empirically test the generic theory under two different ignorance treatments. The first treatment is youth: young firms have less experience than old ones. The second treatment abruptly introduces firms operating in a master-planned economy to a free market—learning to read demand from market signals instead of from state orders takes time. We discuss the results from these empirical tests in the next two subsections respectively.

5.1. Learning with firm age

We estimate the following generic empirical model, which captures not only the effects of age (experience), market volatility and signal noise, but the potential effects of firm size, exporter status, idiosyncratic firm constraints, and industry competitiveness:

$$\text{ErrMag}_{ijm} = \sum_{n=1}^{3} \theta_n \text{Age}_{it}^{n} \tag{10}$$

$$+ \quad \gamma_1 \ \text{SigNoise}_{jm} + \gamma_2 \ \text{MktVol}_{jm} + \alpha_i + \delta_m \tag{11}$$

$$+ \quad \beta_1 \ \ln(\text{Empl}_{\text{it}}) + \beta_2 \ \text{Exp}_{\text{im}} + \beta_3 \ \text{FrmState}_{\text{im}} + \beta_4 \ \text{ImpPen}_{\text{it}} \tag{12}$$

where the forecast ability is measured by the forecast error magnitude (ErrMag). The dependent variable is a cubic function of the firm's age (Age) in (10). Next, we control for signal noise (SigNoise) and market volatility (MktVol) to separate forecast ability from the observed forecast error in (11). We include firm, α_i and calendar month (or month-year) δ_m fixed effects. Lastly, we control for firm attributes: size of the firm (Empl) as measured by the log of production employees, an indicator for whether the firm exports (Exp), an indicator for whether the firm self-reports being constrained (FrmState), and industry competitiveness as measured by import penetration (ImpPen) in (12). The sample is Western firms only, the estimator is OLS, and errors are clustered at the firm level.

Table 2 gives the results. Column (1) concentrates on firm age as an explanatory variable for forecast ability, using (10) and (11). Older firms make lower forecast errors; firms learn as they age. On average, a firm reduces its forecast error by about 9 percent when it is 10 years older. Admittedly, we lack a representative sample of firm ages. In particular, our founding date reflects a sample bias: firm founding dates were collected from recent Amadeus databases. Thus, failed firms, being dropped from the Amadeus data, disproportionately lack founding dates in our sample. Hence, although firm fixed effects ensure that the observed learning does not merely result from the fact that innately better forecasters survived, the estimates of learning are for this rather select subsample, not the population of firms overall. Furthermore, the fact that firms are undersampled in their youth, may, at least partially, explain why the higher order age coefficients are insignificant—the bulk of our variation occurs for mature firms, after learning rates have settled. As predicted by the theory, both signal noise and market volatility are positively related to larger errors, in all models.

Column (2) estimates the effect of firm level attributes besides age, (11) and (12), because age introduces sample bias as discussed above. Larger firms make larger errors. Exporters' forecast errors do not differ significantly from non-exporters'. Firms reporting "constrained" production make much larger errors. Although a number of *ex post* rationalizations are possible, it is not obvious why a factor that the firm is fully aware of, such as self-reported constraints on its

Table 2: Learning over Age

	(1)	(2)	(3)
Firm Age	-0.0061^{***}		-0.0066^{*}
	[0.000]		[0.048]
Firm Age^2	0.0000		-0.0000
	[0.251]		[0.323]
Firm Age^3	-0.0000		0.0000
	[0.106]		[0.290]
Mkt. Volatility	10.9718^{**}	12.7736^{**}	10.6058
	[0.003]	[0.001]	[0.070]
Sig. Noise	5.2614^{***}	4.9000^{***}	5.0184^{***}
	[0.000]	[0.000]	[0.000]
log(Production Employees)		0.0165^{*}	0.0030
		[0.011]	[0.728]
Exporter (=1)		-0.0041	0.0293
		[0.816]	[0.164]
Firm State (Constrained =1)		0.1001^{***}	0.1061^{***}
		[0.000]	[0.000]
Import penetration		0.0646	0.0936
		[0.077]	[0.075]
Constant	-4.2514^{***}	-4.2427^{***}	-3.8130^{***}
	[0.000]	[0.000]	[0.000]
Firm Fixed Effects	*No*	*No*	*No*
Month-Year Fixed Effects	*No*	*No*	*No*
Observations	520310	522638	255997
R^2	0.005	0.006	0.007

p-values in brackets

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: The dependent variable is the squared forecast error. Column (1) includes firm age and market attributes. Column (2) includes firm attributes other than age and market attributes. Column (3) includes all variables. The sample is restricted to Westerns firms only. The estimator is OLS and errors are clustered at the firm level.

production, should necessarily impair forecasting. Firms in industries where import penetration is higher make larger forecast errors but the effect is not statistically significant at the 5 percent level. One might expect import penetration, as a proxy for competitiveness, to increase the incentives to forecast well, but import penetration also proxies for the trade openness of an industry. The positive coefficient suggests that forecasting in open industries is harder.

Column (3) combines all firm attributes, using (10)-(12). The results are similar.

Although measuring firms' learning by examining how forecast quality evolves over their lifespans is intuitive, and we do indeed find evidence that firms learn as they age, detecting more subtle evidence for the other predictions of our theory is limited by our data quality. For example, (omitted) regressions including the second-order effects of age—interactions MktVol $\times$ Age and SigNoise $\times$ Age—yields no significant results. This may be due to the oversampling of older firms whose understanding of the market has largely settled before they enter our sample. Furthermore, firm age may correlate with unobserved firm or market attributes affecting forecast quality. For example, in addition to new firms being smaller (for which we can control), their employees also tend to be younger and have different human capital, and their markets tend to be newer and utilize different technologies. Therefore, in the next section, we look for evidence of firm-level learning after a diverse subsample of firms receive a systemic to shock to their understanding of the market.

5.2. Introduction to the free-market

Relatively homogeneous Germany was abruptly divided in 1949, and for four decades firms in East Germany operated under a masterplanned, communist economy. For these firms of all sizes, maturities, and across the spectrum of industries, market states were dictated, not predicted. Then suddenly, and quite unexpectedly, with German Reunification in 1990, these firms were thrust into the free market economy of the West. Uniquely among transition countries, East Germany immediately received developed country institutions (e.g. legal system, property rights, social welfare) as well as full global market access (Dornbusch et al., 1992). Nevertheless, Eastern managers recognized a deficiency in their understanding of market economies. In 1991 West German firms hosted East German managers as interns. About 70 percent of these interns self-reported having a poor knowledge of market economics; more than 85 percent of their Western hosts shared that assessment (Icks, 1992).

German Reunification offers a natural ignorance shock to East German firms' understanding of the market. However, there is a worry that Reunification left Eastern firms not only with different understandings of the market, but different market conditions altogether, than Western ones. Here we provide evidence that changes in market states did not differ fundamentally between East and West. Differences in forecast errors stem from differences in expectations not realizations.

First, previous research suggests that after Reunification Eastern firms did not sell into different markets. Hitchens et al. (1993, p. 34) show that Eastern firms reoriented their exports from planned to market economies. The reason is twofold. First, demand in transition countries generally collapsed. Second, suddenly these countries had to pay for their imports from former East Germany in Deutschmarks, which they could not afford. The authors show that just before and after Reunification just under 60 percent of Eastern firms' sales were domestic. Sales to former West Germany roughly doubled, while sales to eastern Europe and the former USSR roughly halved in 1991.[9]

Second, our data also indicates that the market states did not differ substantially between the two regions. Figure 3 plots the time series for the Pearson correlation coefficients between Eastern and Western aggregate realizations and expectations respectively (using 7 year rolling windows). The correlation between Eastern and Western aggregate realizations rises rapidly above 0.9 almost immediately after Reunification and increases only slightly thereafter. Correlations between aggregate expectations reach similar strength only after 1997. This suggests that markets between regions homogenized, and the convergence in forecast errors, which we find in our subsequent regressions, does not come from alignment of actual market conditions but rather expectations, which took longer to converge.

[9]This result is based on a survey of 32 firms in the East and 34 firms in the West in 1991. The firms are from the engineering, furniture, clothing, food, and misc. industry categories.

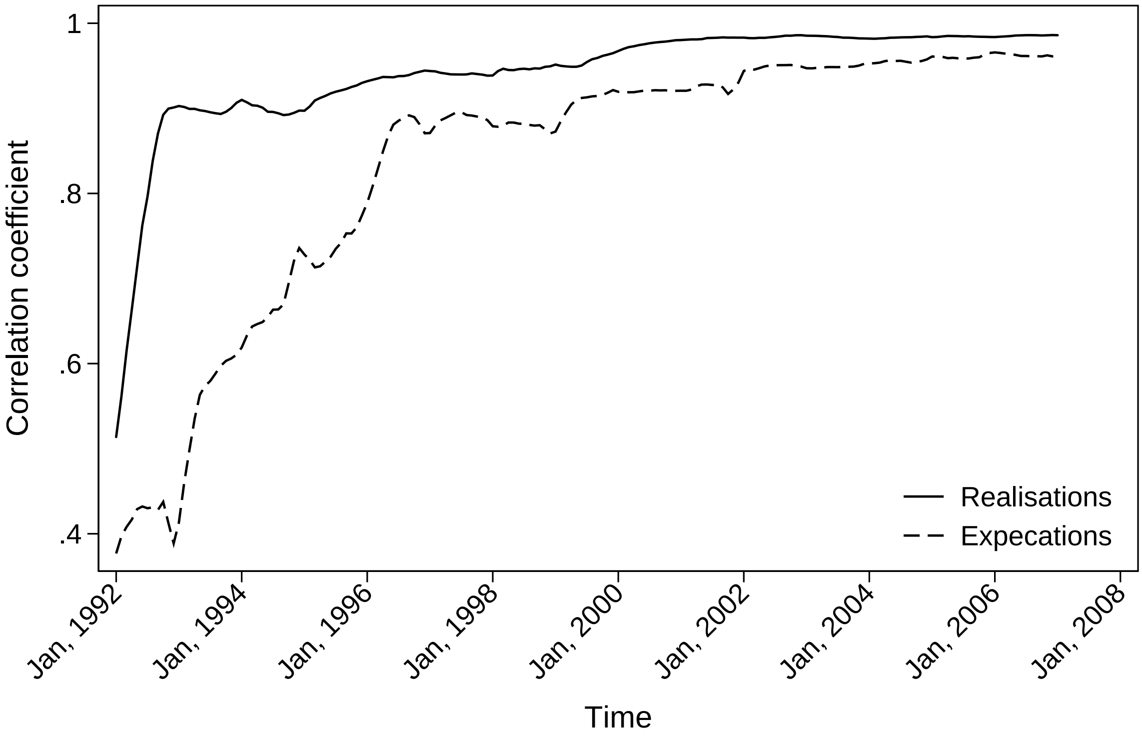


Figure 3: Components of the Market State Forecast Error

Notes: In this graph the lines plots the rolling correlation coefficients (7 year windows) between East and West. The solid line is for aggregate realizations and the dashed line is for aggregate expectations.

To test for learning with experience after Reunification we use the following generic empirical model:

$$\begin{aligned}
\text{ErrMag}_{ijm} &= \beta_1 \text{East}_i + \beta_2\ \text{MktVol}_{jm} + \beta_3\ \text{SigNoise}_{jm} + \alpha_j + \delta_m && (13)\\
&+ \sum_{n=1}^{3} \phi_n \text{t}_m^n \times \text{East}_i && (14)\\
&+ \sum_{n=1}^{3} \delta_n\ \text{MktVol}_{jm} \times \text{t}_m^n \times \text{East}_i && (15)\\
&+ \sum_{n=1}^{3} \rho_n\ \text{SigNoise}_{jm} \times \text{t}_m^n \times \text{East}_i && (16)\\
&+ \gamma_1\ \text{FrmState}_{im} + \gamma_2\ \ln(\text{Empl}_{it}) + \gamma_3\ \text{Exp}_{im} + \gamma_4\ \text{ImpPen}_{it} && (17)\\
&+ \gamma_5 \text{Age}_{it} + \gamma_6 \text{SqAge}_{it}
\end{aligned}$$

The dependent variable is the firm level forecast error magnitude. The first line (13) contains the level effect of East and the first order effects of market volatility (MktVol) and the signal noise (SigNoise). Our theory predicts that the average firm in the East makes larger forecast errors than in the West ($\beta_1 > 0$), as do firms in volatile markets ($\beta_2 > 0$) and those with noisy signals ($\beta_3 > 0$). We also control for industry (α_j) and calendar month (δ_m) fixed effects. The period fixed effects control for common temporal shocks in East and West. That is, identification is in the cross-section within industries and months. The second line (14) includes learning in the East independent of other covariates. Given higher forecast errors in the East initially this implies $\phi_1 < 0$, $\phi_2 > 0$, and $\phi_3 < 0$. Convergence between East and West is modeled using a cubic trend which captures the common experience of firms in the East after Reunification. We do not include a time trend for the West as the theory predicts no significant change in Western errors over time (our empirical counterfactual). Also, we include common period fixed effects δ_m, which would be collinear with a trend for the West. (15) has learning in the East over market volatility. The theory includes the δ_ns but makes no unambiguous prediction for their signs. Learning over signal noise is captured in (16). Our theory predicts convergence over signal noise: $\rho_1 < 0$, $\rho_2 > 0$, and $\rho_3 < 0$. Learning reduces errors faster when signal noise is high. Finally, (17) contains controls: an indicator if the firm self-reports production constrains (FrmState), firm size (Empl), an indicator for whether the firm exports (Exp), import penetration (ImpPen), the firm's age (Age) and age squared (SqAge). As discussed above the inclusion of these control variables (especially age) may introduce sampling bias. And in any case reduces the sample size substantially. Therefore in our preferred specification we replace (17) with a single variable: the square of the actual *ex post* realization of business state reported by the firm. This variable isolates the effects of the firm's predictions while controlling for any idiosyncratic firm shocks.

The monthly trend variable is divided by 12 and normalized to 1992 = 0 to facilitate the interpretation of the estimated coefficients below. The estimator is OLS. For all models, standard errors are clustered at the firm level. We estimate several variations of the generic model in (13)-(17).

Table 3 reports the results for the evolution of forecast errors after the East is treated with a new market state generating process or, alternatively, ignorance. All models control for market volatility and signal noise. As predicted, both increase forecast errors in all models (see eqns. (5) and (6) for predictions). Column (1) estimates the effect of learning over time without independently measuring the dynamic effects of market volatility and signal noise, using (13)-(14). We find broad support for the predictions of the theory: Eastern firms, shocked with market ignorance forecast worse. After Reunification East and West converge, at a diminishing rate (see eqns. (4) and (7)). In Column (2) we estimate the complete learning model, consisting of (13),(15), and (16). We omit (14) because it is nested. The theory's prediction for market volatility's role in the rate of learning is ambiguous (see (9)), and indeed we find no consistent or significant pattern in the volatility-time-east interaction across all models. We find support for the theory's prediction that noise in firm level signals of the future market state positively affects the learning rate (see eqn. (8)). Column (3) adds the controls specified in (17) but results do not change qualitatively. In Column (4) we replace these controls in favor of the actual *ex post* realization of business state reported by the firm. Column (4) is our preferred model, which we use as a benchmark for several robustness tests in the next section.

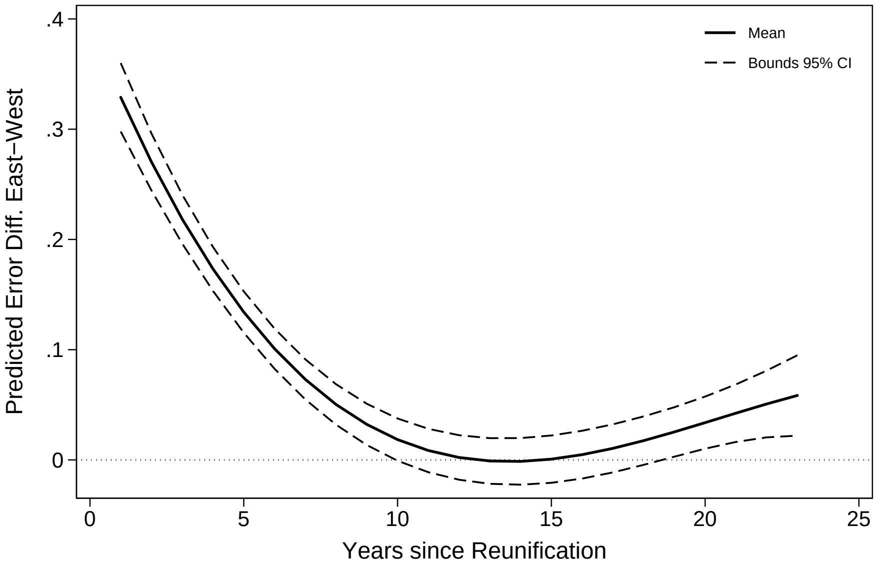
From the coefficients in Column (4) we estimate that it takes approximately 10 years before forecast errors in East and West converge, i.e. cannot be distinguished at a 5 percent level. Figure 4a plots the predicted differences between forecast errors and their 95 percent confidence intervals for years after Reunification. Our theory predicts that if signals within an industry are noisier, errors will initially be larger for inexperienced firms, but learning in such industries reduces forecast errors faster. In our sample, we observe the noisiest signals in the machinery sector. Figure 4b gives the convergence results for this subsample. The speed of convergence (as measured by the slope of the solid lines) is roughly twice the manufacturing average.

Table 3: Learning after Reunification

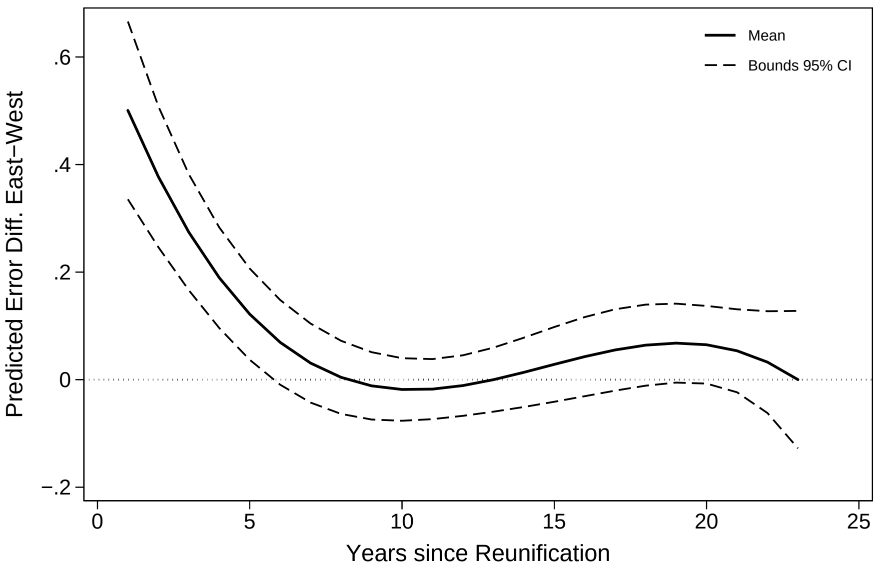
	(1)	(2)	(3)	(4)
East (=1)	0.4048***	0.4071***	0.2516***	0.3934***
	[0.000]	[0.000]	[0.000]	[0.000]
Time × East	−0.0991***			
	[0.000]			
$\text{Time}^2 \times$ East	0.0078***			
	[0.000]			
$\text{Time}^3 \times$ East	−0.0002***			
	[0.000]			
Mkt. Volatility	11.0057***	11.1887***	5.9728	10.0119***
	[0.000]	[0.000]	[0.238]	[0.000]
Mkt. Vol. × Time × East		0.9118	6.3028	0.2155
		[0.504]	[0.082]	[0.864]
Mkt. Vol. × Time^2 × East		−0.0867	−1.1461*	−0.0255
		[0.636]	[0.040]	[0.880]
Mkt. Vol. × Time^3 × East		0.0014	0.0477*	0.0002
		[0.817]	[0.033]	[0.966]
Sig. Noise	3.7684***	3.8479***	3.7029***	3.6100***
	[0.000]	[0.000]	[0.000]	[0.000]
Sig. Noise × Time × East		−0.1052***	−0.0653*	−0.1002***
		[0.000]	[0.016]	[0.000]
Sig. Noise × Time^2 × East		0.0084***	0.0057	0.0081***
		[0.000]	[0.132]	[0.000]
Sig. Noise × Time^3 × East		−0.0002***	−0.0002	−0.0002***
		[0.000]	[0.289]	[0.000]
State Realisation^2				0.4145***
				[0.000]
Firm State (Constraint =1)			0.1421***	
			[0.000]	
log(Prod. Employees)			0.0036	
			[0.362]	
Exporter (=1)			0.0072	
			[0.601]	
Import penetration			0.0651	
			[0.111]	
Firm Age			0.0003	
			[0.317]	
Firm Age^2			−0.0000*	
			[0.035]	
Constant	−3.1801***	−3.2200***	−3.0191***	−3.0805***
	[0.000]	[0.000]	[0.001]	[0.000]
Industry Fixed Effects	*Yes*	*Yes*	*Yes*	*Yes*
Month-Year Fixed Effects	*Yes*	*Yes*	*Yes*	*Yes*
Observations	1446809	1446809	312832	1446809
R^2	0.008	0.008	0.013	0.038

p-values in brackets
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: (1) learning is only over time. (2) learning is over volatility and signal noise. (3) adds observed firm level attributes. (4) substitutes the realized market state magnitude for firm level attributes. The estimator is OLS and errors are clustered at the firm level.



(a) All manufacturing



(b) Machinery

Figure 4: Convergence

Notes: This graph plots the predicted difference in forecast error magnitude between East and West and its 95% confidence interval.

6. Robustness

Table 4 contains the results of several robustness tests. Column (1) is the same model as Column (4) in Table 3 for reference. Despite the inclusion of industry fixed effects there might be relevant, unobserved firm heterogeneity, and learning could occur both within the firm and at market level due to better survival rates for good forecasters. In Column (2) we replace industry fixed effects by firm level fixed effects (dropping the time-invariant East indicator). The effects for learning over noise are qualitatively similar but smaller as they might be correlated with unobserved heterogeneity. Our estimates with fixed effects do not include any learning that may be due to poorer forecasting firms exiting the market. Nevertheless, learning could be identified by poor forecasters that exit shortly after Reunification. Therefore, in Column (3) we restrict our sample in the East to firms that first answered the survey in 1992 and answered at least once more after 1999, which eliminates 75 percent of Eastern firms. Since the West is in steady state, we do not similarly restrict the Western sample. Quantitatively, the learning effects are similar to the effects in Column (2), providing more evidence that learning is not primarily driven by exit. Our theory describes changes in mean squared error, but to ensure that our results are not an artifact of the magnification of large errors through squaring, in Column (4) we estimate a model using absolute value instead, yielding the same qualitative results. Fourth, since as we discussed above, non-neutral forecasts correlate with firm attributes, we restrict our sample to non-neutral forecasts in Column (5) to confirm robustness to unknown biases related to neutral forecasts. Qualitatively, the results are the same but all the effects are larger. Firms in the East initially make larger errors and the learning rate is higher.

Table 4: Robustness

	(1)	(2)	(3)	(4)	(5)
East (=1)	0.3936***	(*dropped*)	0.2970***	0.2096***	0.7414***
	[0.000]		[0.000]	[0.000]	[0.000]
Mkt. Volatility	9.9311***	7.7848***	10.3320***	4.1711***	3.4538
	[0.000]	[0.000]	[0.000]	[0.000]	[0.341]
Mkt. Vol. × Time × East	0.3361	2.9207*	−1.1695	0.0652	1.1174
	[0.791]	[0.027]	[0.447]	[0.934]	[0.725]
Mkt. Vol. × Time^2 × East	−0.0393	−0.2210	0.1810	0.0090	−0.0123
	[0.818]	[0.190]	[0.417]	[0.935]	[0.977]
Mkt. Vol. × Time^3 × East	0.0006	0.0046	−0.0066	−0.0009	−0.0035
	[0.912]	[0.396]	[0.392]	[0.807]	[0.800]
Sig. Noise	3.5646***	3.8496***	3.7859***	1.5461***	1.7098
	[0.000]	[0.000]	[0.000]	[0.000]	[0.066]
Sig. Noise × Time × East	−0.1006***	−0.0785***	−0.0726***	−0.0512***	−0.2201***
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Sig. Noise × Time^2 × East	0.0081***	0.0058***	0.0054**	0.0040***	0.0174***
	[0.000]	[0.000]	[0.003]	[0.000]	[0.000]
Sig. Noise × Time^3 × East	−0.0002***	−0.0001**	−0.0001*	−0.0001***	−0.0004***
	[0.000]	[0.001]	[0.049]	[0.000]	[0.000]
State Realisation^2	0.4145***	0.3752***	0.4141***	0.3240***	−0.5711***
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Constant	−3.0465	−3.2374***	−3.2476	−1.0905	−0.4016
	[0.879]	[0.000]	[0.697]	[.]	[0.678]
Firm Fixed Effects	*No*	*Yes*	*No*	*No*	*No*
Industry Fixed Effects	*Yes*	*No*	*Yes*	*Yes*	*Yes*
Month-Year Fixed Effects	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Observations	1446794	1446794	1357463	1446794	520782
R^2	0.038	0.024	0.037	0.052	0.044

p-values in brackets

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: (1) is the same as column (5) in Table 3, (2) substitutes firm fixed effects for industry fixed effects, (3) restricts the Eastern sample to firms that survived the period 1992 to 1999, (4) uses the absolute error as the dependent variables, (5) restricts the sample to non-neutral expectations.

The forecast error calculation in section 4 is only one of many possible definitions. Above we tested the robustness of our results to the absolute value of error, rather than squared error. In section B in the appendix we also investigate the robustness of our results to different definitions of the forecast error itself as well as the use of a nonlinear probability model. The economic and statistical significance of individual coefficient estimates are largely robust to these alternative specifications. The estimated time for convergence between Eastern and Western error size varies by less than a year in all specifications.

7. Conclusion

Macroeconomic theory hinges on market actors being able to predict future market characteristics. For nearly forty years, learning has been offered as a theoretical justification for the dominant paradigm to describe expectation formation: *rational expectations*. In this pattern, we also introduced a formal model of Bayesian learning, in which firms learn the distribution of market states, and thereby improve their forecasting ability. Unlike previous models, which prove equilibrium attainment, ours focuses on how several distinct types of uncertainty influence the quality and rate of improvement in forecasts. This theory predicts that firm's forecasts improve with experience, but at an ever decreasing rate. The model predicts that both market volatility and signal noise make forecasting more difficult but that firms in noisier markets reduce their forecast errors faster.

We find that the theory's first-order predictions over experience, when firm age is used as a proxy, are borne out in firm-level forecasting data. However due to the correlation of age to a number of other unobserved firm features that could plausibly explain forecast error differences, and due to peculiar limitations in our data, we also conduct an alternative evaluation of the theory. By comparing firms in former East and West Germany that survived the Reunification of Germany, we can test whether Eastern firms learn how to predict market states. They do. When time from Reunification proxies for experience, we find empirical support for all of the theory's predictions. Of particular importance, forecast quality between Eastern and Western firms converges...after a decade. Our evidence suggests that this delay is not due to slow convergence of the markets themselves, as these align quickly but due to gradual improvement in predictions by Eastern firms.

Although the patterns we observe in the data are consistent with the theory, the mechanisms by which firms learn remain opaque. Neither the theory nor our empirics can distinguish between institutional learning and the learning of individual Eastern managers. Although we have ruled out survival of the fittest at the firm level as primary driver of the observed improvements, we cannot rule out that better forecasting managers (perhaps Western ones) displace worse ones within firms.

Finally, although we have shown that the magnitude of Eastern firms' forecast errors converge to those of their Western counterparts, we have not shown that those Western firms' expectations are *rational.* In fact, a t-test rejects the null hypothesis that Western forecast errors over our entire sample period have zero mean with greater than 99.9% confidence. They are systematically overoptimistic—they predict that the future state of business will be better than they actually report that it is when the time comes.[10] Although we know of no economic theory to explain systematic biases in firms' expectation formation, optimism bias has been consistently documented in healthy individuals. The consequences of such potential biases are subjects for future research.

[10]Bachmann and Elstner (2015) also show that a majority of German firms are over overoptimistic.

References

Alesina, A. and Fuchs-Schündeln, N. (2007). Good-Bye Lenin (or Not?): The Effect of Communism on People's Preferences. *The American Economic Review*, 97(4):1507–1528.

Bachmann, R. and Elstner, S. (2015). Firm optimism and pessimism. *European Economic Review*, 79:297–325.

Bachmann, R., Elstner, S., and Sims, E. R. (2013). Uncertainty and Economic Activity: Evidence from Business Survey Data. *American Economic Journal: Macroeconomics*, 5(2):217–49.

Bloom, N. (2014). Fluctuations in Uncertainty. *Journal of Economic Perspectives*, 28(2):153–176.

Cho, D. W. and Hersch, P. L. (1998). Forecaster characteristics and forecast outcomes. *Journal of Economics and Business*, 50(1):39–48.

Comin, D. and Mulani, S. (2006). Diverging trends in aggregate and firm volatility. *The Review of Economics and Statistics*, 88(2):374–383.

Cyert, R. M. and DeGroot, M. H. (1974). Rational Expectations and Bayesian Analysis. *Journal of Political Economy*, 82(3):521–536.

Doraszelski, U., Lewis, G., and Pakes, A. (2014). Just starting out: Learning and price competition in a new market.

Dornbusch, R., Wolf, H., and Alexander, L. (1992). Economic transition in eastern Germany. *Brookings papers on economic activity*, 1992(1):235–272.

Fritsch, M. and Mallok, J. (1998). Surviving the transition: the process of adaptation of small and medium-sized firms in East Germany. *The Microeconomics of Transition and Growth. Cheltenham: Elgar*, pages 163–184.

Fudenberg, D. and Levine, D. K. (1998). *The theory of learning in games*, volume 2. MIT press.

Heineck, G. and Süssmuth, B. (2010). A Different Look at Lenin's Legacy: Trust, Risk, Fairness and Cooperativeness in the Two Germanies. SSRN Scholarly Paper ID 1688271, Social Science Research Network, Rochester, NY.

Hitchens, D. M., Wagner, K., and Birnie, J. E. (1993). *East German productivity and the transition to the market economy: comparisons with West Germany and Northern Ireland.* Avebury.

Icks, A. (1992). *Mittelständische Unternehmen als Qualifizierungspaten: Betriebspraktika für ostdeutsche Fach-und Führungskräfte; Begleitforschung zu einer Initiative der Stiftung Industrieforschung.* Schäffer-Poeschel.

Kim, Y., Lobo, G. J., and Song, M. (2011). Analyst characteristics, timing of forecast revisions, and analyst forecasting ability. *Journal of Banking & Finance*, 35(8):2158–2168.

Muth, J. F. (1961). Rational Expectations and the Theory of Price Movements. *Econometrica*, 29(3):315–335.

Nerlove, M. (1958). Adaptive Expectations and Cobweb Phenomena. *The Quarterly Journal of Economics*, 72(2):227–240.

Nerlove, M. (1983). Expectations, Plans, and Realizations in Theory and Practice. *Econometrica*, 51(5):1251–1279.

Turnovsky, S. J. (1969). A bayesian approach to the theory of expectations. *Journal of Economic Theory*, 1(2):220–227.

van Hoorn, A. and Maseland, R. (2010). Cultural differences between East and West Germany after 1991: Communist values versus economic performance? *Journal of Economic Behavior & Organization*, 76(3):791–804.

Vives, X. (1993). How Fast do Rational Agents Learn? *The Review of Economic Studies*, 60(2):329–347.

A. Neutral expectations

Table 5: Neutral expectations

	(1)
East	-0.0403
	[0.432]
Time	0.0002^{***}
	[0.000]
East $\times$ Time	-0.0001
	[0.338]
Firm Age	0.0004^{***}
	[0.000]
Response freq.	0.0001^{***}
	[0.000]
log(Production Employees)	0.0117^{***}
	[0.000]
Exporter (=1)	-0.0114
	[0.064]
Firm State (Constrained =1)	-0.2903^{***}
	[0.000]
Import penetration	-0.0378^{***}
	[0.000]
Market Volatility	-86.4346^{***}
	[0.000]
Sig. Noise	-35.7390^{***}
	[0.000]
Constant	35.9426^{***}
	[0.000]
Observations	339684

p-values in brackets
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: The dependent variable is an indicator that takes the value 1 if the firm gives a neutral expectation for the market state. Probit regression.

B. Different Error Definitions

In this section we test the robustness of our results to different definitions of the forecast errors. First, we eliminate the distinction between small and large errors, that is we only differentiate between error (1) or no error (0). The results are in Table 6. Given this binary error definition we also investigate the robustness of our results to the use of a probit estimator. The coefficient estimates are in Table 7. Second, we maintain the distinction between small and large errors but change the parametrization as introduced in section 4. Recall that the specifications in the main text use the following parameterization when calculating the errors: $H^h = \frac{2}{3}$, $H^m = \frac{1}{3}$ and $H^l = 0$. To test robustness to this parameterization we make the requirements to obtain an error more or less stringent. We make it harder to obtain a (large) error by defining: $H^h = \frac{3}{4}$, $H^m = \frac{1}{3}$ and $H^l = -\frac{1}{3}$. Table 8 gives the estimation results. And we can make it easier by defining: $H^h = \frac{1}{2}$, $H^m = \frac{1}{4}$ and $H^l = 0$. Table 9 gives the results. Across these robustness tests the magnitude of the coefficient estimates naturally varies. For instance, when it is harder (easier) to make large forecast errors the coefficient estimates for the level effects are lower (higher). Importantly, the estimated convergence date between East and West does not vary by more than by one year. Like in the main text we used the models from Columns (4) to predict the forecast errors and plot the differences between East and West in Figure 5.

Table 6: Learning and Reunification (binary forecast errors)

	(1)	(2)	(3)	(4)
East (=1)	0.1264^{***}	0.1268^{***}	0.0872^{***}	0.1176^{***}
	[0.000]	[0.000]	[0.000]	[0.000]
Time × East	-0.0279^{***}			
	[0.000]			
$\text{Time}^2\times$ East	0.0021^{***}			
	[0.000]			
$\text{Time}^3\times$ East	-0.0000^{**}			
	[0.001]			
Mkt. Volatility	2.0820^{*}	2.1080^{*}	-0.2013	1.3122
	[0.037]	[0.036]	[0.945]	[0.159]
Mkt. Vol. × Time × East		0.3630	3.6876^{*}	-0.1107
		[0.602]	[0.038]	[0.860]
Mkt. Vol. × Time^2 × East		-0.0020	-0.6842^{*}	0.0396
		[0.985]	[0.018]	[0.667]
Mkt. Vol. × Time^3 × East		-0.0011	0.0282^{*}	-0.0019
		[0.756]	[0.016]	[0.548]
Sig. Noise	0.6597^{**}	0.6862^{**}	1.0117^{*}	0.5369^{*}
	[0.007]	[0.005]	[0.040]	[0.015]
Sig. Noise × Time × East		-0.0297^{***}	-0.0176	-0.0263^{***}
		[0.000]	[0.166]	[0.000]
Sig. Noise × Time^2 × East		0.0021^{**}	0.0012	0.0019^{**}
		[0.002]	[0.509]	[0.002]
Sig. Noise × Time^3 × East		-0.0000^{*}	-0.0000	-0.0000^{*}
		[0.048]	[0.761]	[0.037]
State Realisation^2				0.2787^{***}
				[0.000]
Firm State (Constraint =1)			0.0633^{***}	
			[0.000]	
log(Prod. Employees)			0.0045^{*}	
			[0.048]	
Exporter (=1)			0.0017	
			[0.835]	
Import penetration			0.0626^{**}	
			[0.006]	
Firm Age			0.0002	
			[0.191]	
Firm Age^2			-0.0000^{*}	
			[0.016]	
Constant	-0.1408	-0.1514	-0.4338	-0.0673
	[0.574]	[0.547]	[0.382]	[0.766]
Industry Fixed Effects	*Yes*	*Yes*	*Yes*	*Yes*
Month-Year Fixed Effects	*Yes*	*Yes*	*Yes*	*Yes*
Observations	1446820	1446820	312832	1446820
R^2	0.006	0.006	0.010	0.053

p-values in brackets

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: (1) learning is only over time, (2) learning is over volatility and signal noise, (3) adds observed firm level attributes, (4) substitutes the realized market state magnitude for firm level attributes. The estimator is OLS and errors are clustered at the firm level.

Table 7: Learning and Reunification (probit model)

	(1)	(2)	(3)	(4)
Sq. Error				
East (=1)	0.3218***	0.3228***	0.2242***	0.3144***
	[0.000]	[0.000]	[0.000]	[0.000]
Time × East	−0.0714***			
	[0.000]			
$\text{Time}^2 \times$ East	0.0054***			
	[0.000]			
$\text{Time}^3 \times$ East	−0.0001***			
	[0.001]			
Mkt. Volatility	5.1702*	5.2642*	0.5041	3.2336
	[0.045]	[0.042]	[0.947]	[0.196]
Mkt. Vol. × Time × East		0.9528	9.3567*	−0.2649
		[0.590]	[0.040]	[0.874]
Mkt. Vol. × Time^2 × East		−0.0145	−1.7440*	0.0999
		[0.956]	[0.019]	[0.682]
Mkt. Vol. × Time^3 × East		−0.0024	0.0719*	−0.0048
		[0.789]	[0.018]	[0.563]
Sig. Noise	1.7492**	1.8237**	2.9628*	1.4331*
	[0.006]	[0.005]	[0.022]	[0.019]
Sig. Noise × Time × East		−0.0762***	−0.0460	−0.0713***
		[0.000]	[0.155]	[0.000]
Sig. Noise × Time^2 × East		0.0055**	0.0033	0.0052**
		[0.002]	[0.484]	[0.002]
Sig. Noise × Time^3 × East		−0.0001*	−0.0001	−0.0001*
		[0.041]	[0.727]	[0.032]
State Realisation^2				0.7170***
				[0.000]
Firm State (Constraint =1)			0.1595***	
			[0.000]	
log(Prod. Employees)			0.0116*	
			[0.041]	
Exporter (=1)			0.0045	
			[0.829]	
Import penetration			0.1570**	
			[0.006]	
Firm Age			0.0006	
			[0.173]	
Firm Age^2			−0.0000*	
			[0.015]	
Constant	−1.6919*	−1.7253**	−2.7676*	−1.4970*
	[0.010]	[0.009]	[0.034]	[0.016]
Industry Fixed Effects	*Yes*	*Yes*	*Yes*	*Yes*
Month-Year Fixed Effects	*Yes*	*Yes*	*Yes*	*Yes*
Observations	1446809	1446809	312832	1446809

p-values in brackets

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: (1) learning is only over time, (2) learning is over volatility and signal noise, (3) adds observed firm level attributes, (4) substitutes the realized market state magnitude for firm level attributes. The estimator is Probit.

Table 8: Learning and Reunification (error function parametrization: stringent error thresholds)

	(1)	(2)	(3)	(4)
East (=1)	0.3875***	0.3897***	0.2485***	0.3793***
	[0.000]	[0.000]	[0.000]	[0.000]
Time × East	−0.0921***			
	[0.000]			
$\text{Time}^2 \times$ East	0.0072***			
	[0.000]			
$\text{Time}^3 \times$ East	−0.0002***			
	[0.000]			
Mkt. Volatility	12.1171***	12.3115***	8.4261	11.4175***
	[0.000]	[0.000]	[0.089]	[0.000]
Mkt. Vol. × Time × East		0.4198	6.0921	−0.1091
		[0.749]	[0.077]	[0.930]
Mkt. Vol. × Time^2 × East		−0.0537	−1.2024*	−0.0073
		[0.762]	[0.027]	[0.965]
Mkt. Vol. × Time^3 × East		0.0011	0.0526*	0.0002
		[0.852]	[0.017]	[0.970]
Sig. Noise	4.1836***	4.2453***	4.5321***	4.0646***
	[0.000]	[0.000]	[0.000]	[0.000]
Sig. Noise × Time × East		−0.0961***	−0.0683*	−0.0923***
		[0.000]	[0.010]	[0.000]
Sig. Noise × Time^2 × East		0.0077***	0.0070	0.0074***
		[0.000]	[0.064]	[0.000]
Sig. Noise × Time^3 × East		−0.0002***	−0.0002	−0.0002***
		[0.000]	[0.123]	[0.000]
State Realisation^2				0.3148***
				[0.000]
Firm State (Constraint =1)			0.1448***	
			[0.000]	
log(Prod. Employees)			−0.0009	
			[0.807]	
Exporter (=1)			0.0132	
			[0.333]	
Import penetration			0.0655	
			[0.095]	
Firm Age			0.0004	
			[0.174]	
Firm Age^2			−0.0000**	
			[0.009]	
Constant	−3.5993***	−3.6279***	−3.8419***	−3.5219***
	[0.000]	[0.000]	[0.000]	[0.000]
Industry Fixed Effects	*Yes*	*Yes*	*Yes*	*Yes*
Month-Year Fixed Effects	*No*	*No*	*No*	*No*
Observations	1446809	1446809	312832	1446809
R^2	0.007	0.007	0.011	0.023

p-values in brackets

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: (1) learning is only over time, (2) learning is over volatility and signal noise, (3) adds observed firm level attributes, (4) substitutes the realized market state magnitude for firm level attributes. The estimator is OLS and errors are clustered at the firm level.

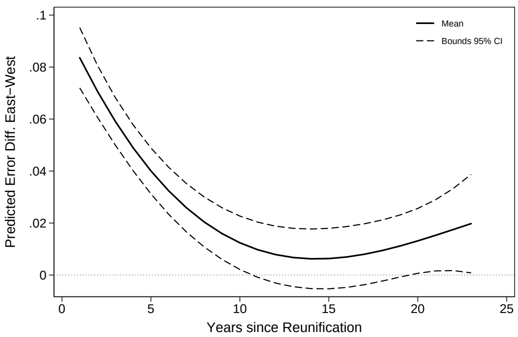
Table 9: Learning and Reunification (error function parametrization: relaxed error thresholds)

	(1)	(2)	(3)	(4)
East (=1)	0.4276^{***}	0.4296^{***}	0.2540^{***}	0.4167^{***}
	[0.000]	[0.000]	[0.000]	[0.000]
Time × East	-0.1022^{***}			
	[0.000]			
$\text{Time}^2 \times$ East	0.0080^{***}			
	[0.000]			
$\text{Time}^3 \times$ East	-0.0002^{***}			
	[0.000]			
Mkt. Volatility	10.6364^{***}	10.8131^{***}	4.2326	9.6989^{***}
	[0.000]	[0.000]	[0.430]	[0.000]
Mkt. Vol. × Time × East		0.7598	6.9087	0.1005
		[0.597]	[0.086]	[0.940]
Mkt. Vol. × Time^2 × East		-0.0571	-1.2533^{*}	0.0008
		[0.770]	[0.042]	[0.996]
Mkt. Vol. × Time^3 × East		0.0003	0.0522^{*}	-0.0008
		[0.964]	[0.034]	[0.893]
Sig. Noise	3.5781^{***}	3.6559^{***}	3.4265^{***}	3.4307^{***}
	[0.000]	[0.000]	[0.000]	[0.000]
Sig. Noise × Time × East		-0.1074^{***}	-0.0598^{*}	-0.1027^{***}
		[0.000]	[0.040]	[0.000]
Sig. Noise × Time^2 × East		0.0084^{***}	0.0049	0.0081^{***}
		[0.000]	[0.232]	[0.000]
Sig. Noise × Time^3 × East		-0.0002^{***}	-0.0001	-0.0002^{***}
		[0.000]	[0.424]	[0.000]
State Realisation^2				0.3924^{***}
				[0.000]
Firm State (Constraint =1)			0.1500^{***}	
			[0.000]	
log(Prod. Employees)			0.0033	
			[0.442]	
Exporter (=1)			0.0055	
			[0.712]	
Import penetration			0.0527	
			[0.220]	
Firm Age			0.0003	
			[0.334]	
Firm Age^2			-0.0000^{*}	
			[0.038]	
Constant	-2.9120^{***}	-2.9512^{***}	-2.7114^{**}	-2.8190^{***}
	[0.000]	[0.000]	[0.003]	[0.000]
Industry Fixed Effects	*Yes*	*Yes*	*Yes*	*Yes*
Month-Year Fixed Effects	*Yes*	*Yes*	*Yes*	*Yes*
Observations	1446809	1446809	312832	1446809
R^2	0.009	0.009	0.013	0.034

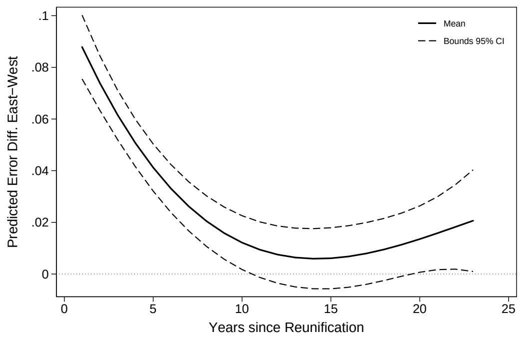
p-values in brackets

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

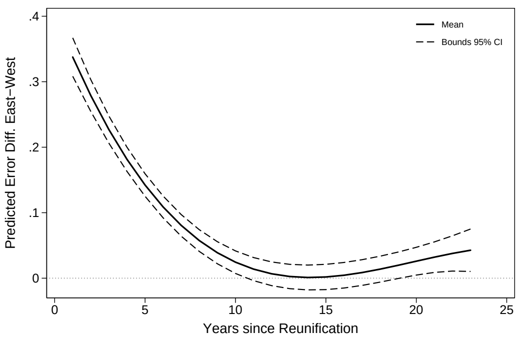
Notes: (1) learning is only over time, (2) learning is over volatility and signal noise, (3) adds observed firm level attributes, (4) substitutes the realized market state magnitude for firm level attributes. The estimator is OLS and errors are clustered at the firm level.



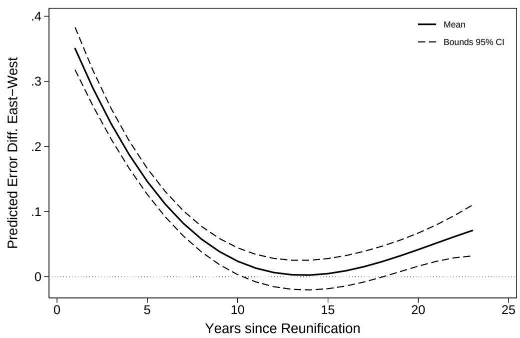
(a) Binary error (OLS)



(b) Binary error (Probit)



(c) Stringent error thresholds



(d) Relaxed error thresholds

Figure 5: Convergence

Notes: This graph plots the predicted difference in forecast error magnitude between East and West and its 95% confidence interval. The predicted values are from the fourth Columns of Tables 6, 7, 8, and 9, respectively.