Referee report
on Harald Tauchmann
"West-East Convergence in the Prevalence of Illicit Drugs: Socioeconomics or Culture"

Summary
This paper uses survey data on illicit drug use since reunification to evaluate the roles of observable characteristics and unobserved "culture" effects in the closing of the East/West illicit drug use gap from 1990-2000.

Comments
1. The main data seem extremely limited in terms of comparability and available drug use measures. The fact that question wording changed multiple times over the course of the sample period is very troubling, since there is a wealth of research showing that how one asks questions about substance use matters quite a bit for the reported rates. At a minimum, an appendix with actual question wording for the main variables should be provided, particularly since readers will not be very familiar with these data. 2. Probably the most problematic aspect of the drug use measure is that it groups so many different measures of drug use into a single illicit drug use measure. Cannabis is a decidedly different substance than is cocaine (and LSD, amphetamines, etc.). These substances can differ dramatically in price, availability, prevalence, purity, effects on health, and other important characteristics. This means that aggregating all the substances together into an "any illicit drug use" outcome is completely inappropriate. 3. Not using the information on frequency of use – i.e. restricting attention only to past year use – also seems very problematic. The problem is that the author is throwing away potentially important and useful information. While I appreciate the basic intuition for why past year use might be the construct of interest for this paper, it follows that, say, lifetime use might be a useful "control" outcome for evaluating the socioeconomics/culture question. I recommend all available data be used to estimate multiple models, which would allow for a systematic analysis of patterns across different use measures and substances. 4. The lack of availability measures is a major problem for this analysis, since the observed rates conflate supply and demand factors. This has very important implications for evaluating the main socioeconomics/culture claim in the paper, since systematic region/time variation in availability will bias the observed convergence estimates. 5. Presumably these surveys also include information on non-illicit substances like alcohol and tobacco. This may be a way to address the problem of lack of information on availability measures, in part because there should be more variation: availability differs dramatically across substances. I recommend similar analyses of tobacco and alcohol use since presumably there should be much cleaner availability measures (prices, taxes, liquor outlets, smoking laws, etc.). It would also be an important comparison to the drug use outcomes for evaluating the socioeconomics/culture question, since I imagine (but do not know) preexisting differences in smoking and drinking were much more similar than for illicit drug use. Showing whether these substances similarly converged or diverged would help the reader better evaluate the main socioeconomics/culture claim in the paper. 6. Footnote 6 indicates that other variables were assessed and available (e.g. body height and weight) but were not included in the analysis. This seems strange to me, particularly since this paper is not about using a design to isolate purely exogenous variation in right hand side variables. If the author is not going to worry about controlling for clearly endogenously determined covariates in the baseline

specification – education, marital status, employment, income, and location are all codetermined with substance use – then it makes sense to put in all available observed covariates (i.e. "kitchen sink" the model). Directions of causality to/from drug use and physical health (e.g. weight/BMI) are not obvious, but since this is true for the other covariates as well the model should include these and other consistently measured variables. 7. What do the survey weights weight for? That is, the weights are designed to make the survey representative at what level? 8. Some discussion of the likely effects of survey mode is warranted. Since people are reporting illicit behavior, systematic underreporting is a concern. The fact that these surveys were mail based generally increases the importance of these concerns, since individuals likely infer that the surveyors can match their survey to their address, such that reporting illicit behavior will be even more underreported. There is literature on this in the substance abuse area that should be referenced and thought about for this analysis. 9. The elimination of family background characteristics after 1995 is a serious problem, since family background controls are likely some of the strongest predictors of illicit drug use (at least this is the case for US analyses). It would be useful to estimate models on the subsample of surveys that include this information to get a sense of how far off the later models are that do not include this information. 10. The construction of the sample is meant to effectively account for time effects/convergence in illicit drug use. An alternative approach would be to use the younger individuals in the earlier surveys and create synthetic cohorts across the later survey waves, as is common in health and labor contexts.