Points to remember

* In the identification section
  + Rho is the correlation between Z2 and the true expectation
    - In our main model, we don’t think of Z2 as a second measurement that is a function of X\* and an error. Instead, all we require is the E[e|X\*,Z2,v]=0, so the expectational error is mean independent of the instrument and the structural error (of course, Z2 also has to be sufficiently correlated with X2 or else the set might be informative/not bounded)
    - In the stricter model that is part of the proof, we show that a model that does not impose corr(X\*,e\_Z) will only set identify sigma\_X\*, and so only set identify beta. One needs a second measurement of X\* in order to point identify beta. Even in our stricter model with normality assumptions on the errors, etc, we’re not point identified
  + In this identification section, it doesn’t make sense economically to call Z2 past period’s revenue. That only makes sense when we’re thinking of our actual statistical model in which the only condition we impose on the instrument is that it’s mean independent of the expectational error and fully independent of the structural error (so that there’s no endogeneity arising from the structural error). In our statistical model, it makes sense to think of the Z2 as past period’s revenue, because if it’s in the info set of the agent, it will be independent of the contemporaneous expectational error. It will be independent of the structural error is you believe the error in profit function is not correlated over time (maybe a little harder to believe, but we can always use multiple period lagged revenue to make the lack of correlation in the unobservables more believable).
* Score function
  + The idea is to use information on the observed binary choices to define the log probability of the observed choices. Taking the first derivative wrt beta, get a score function that is set equal to zero as a function of the true X\* (the unobserved expectations of the agent). When we put in a proxy for revenue (X\_hat in place of X\*), we get an expectational error. Assumption 2 gives that the form of the structural error guarantees that the ratio of cdfs in the score function will be convex, such that a change in the truncation point to include the expectational error will turn the score function equality to an inequality in the direction that makes the inequality weaker (so that it will contain the truth; just need Jensen’s inequality and the convexity in the truncation point for this to hold).
  + This is not using much more information than the revealed preference inequalities, at least in the case in which we assume there is a structural error and therefore need to account for it in the revealed preference inequalities. One difference though is that the revealed preference inequalities condition on the decision, which means you have to deal with selection on the unobservable when forming the conditional expectation. The score function moments don’t condition on the decision.
  + The one case in which we know something about sharpness is when there is no measurement error and we know there is none (meaning we can use the observed X as the “instrument”). In that case, the score function moments will return a point, equal to the true point, while the revealed preference moments will return a set even without measurement error. This is the only case for which we know for sure that the revealed preference moments are redundant.
* “Truth” for counterfactuals
  + We recover the parameter estimates absent bias from the presence of measurement error. However, one problem we face in predictions is that is we apply X\_obs against our estimates, we’d need to know something about the distribution of the measurement error in order to make predictions.
  + This limits us to performing counterfactuals that involve the social planner or firm specifying the expectations the agent will use for their entry decision. For example, the export authorities might provide exporters a guaranteed nominal exchange, which would remove the expectational error in their forecasts. We could test a counterfactual environment in which such an environment changes. We’re still thinking about what kind of richer counterfactuals might be possible.
* Is the approach “weak IV robust”
  + Depends what weak IV means. This is not an identification question (identification doesn’t depend on the strength of the instruments, persay). It is an estimation question. In our approach, if the correlation between Z and X is small, we’ll find inequalities that are wide and possibly unbounded. It’s robust in the sense that it wouldn’t give us a misleadingly precise finding about the values of our parameters.
* Fail to find a set that satisfies all of the inequalities. Is this due to sampling error or a misspecification that neglects structural error?
  + Can test if the inequalities are “true” using Bugni, Canay, Shi and others. If the equalities fail the test, then it’s not sampling error but model misspecification. Would be important to know whether, if fail to get a set, we can blame it on sampling error vs. misspecification
    - In out simulation exercise, we know it’s not due to sampling error because we use the same population that generated the truth??
* When running a probit (which is what we’d do if we assumed no measurement error), aren’t you leaving off information in Z?
  + Well, on, because there’s no need to include Z when the model assumes no measurement error.
  + Also, in general, in this case our model imposes fewer assumptions. It’s probably better to use a model with fewer restrictions unless you know for sure that the assumptions are guaranteed to hold. Here, our approach is robust to expectational error, probit is not.
* What is the extra info in the score function moments over the RP moments?
  + Hard to say what’s sharper, outside of the specific setting mentioned above (when we know there is no expectational error; in that case RP moments are redundant).
  + Score function uses information on the form of the log probability of a choice that depends on the structural error, nu. When nu is distributed normally, the score function involves a ratio of CDFs that can be calculated simply in Matlab. When there is expectational error, we rely on Jensen’s inequality along with the convexity of the ratio in the truncation point to get a inequality between the ratio involving measurement error and the ratio without measurement error. Substituting in for the equality in the expression, we get an inequality that will always be weaker than the equality that comes in the case of no measurement error (or if measurement error increases, the inequality will also simply become larger but will preserve the direction of the inequality).
  + Score function moments not conditional on choice, just conditioning on X and v distribution; revealed preference inequalities condition on the choice. When we assume there is no v, then the distribution of the measurement or expectational error doesn't matter, because the choice is independent of v. However, when there is a v, need to be careful to calculate the conditional expectation of v conditional on the choice, since this will truncate the distribution of the structural error conditional on the choice.