Specification of Literature and Default Priors

Our registered report notes that in addition to using priors elicited from experts, policymakers, and the firms themselves, we will also estimate results using literature-informed priors and default priors. This document provides more detail on how we will form these literature and default priors.

Literature Priors

The modern procedure for Bayesian data analysis advocated in Gelman et al (2004) suggests the use of previous literature to inform priors in the analysis of any new study data. In this document we record the literature priors for the 10 key parameters of interest related to our causal inference questions:

1. *Take-up rate among those offered treatment*: Gaussian with truncated support at 0 and 100, with mean 80% and standard deviation 21%. The mean is specified in our registered report, based on McKenzie and Woodruff (2014) reporting an average take-up rate of 65% for firms in business training experiments, adjusted upwards to reflect firms self-selected into the program. The 80% rate also falls midway between the 87% take-up rate for individual consulting and 75% take-up rate for group consulting reported for a management consulting program in Colombia by Iacovone, Maloney and McKenzie (2019). The standard deviation of 21% implies a truncated 95% confidence interval of (39%, 100%), which includes the take-up rates for all business training experiments reported in McKenzie and Woodruff (2014).

Our approach for literature-informed priors for the next six outcomes follows the approach used in our registered report for power calculations. It is based on the finding of McKenzie and Woodruff (2017) that an approximate estimate of the treatment effect of a business training intervention on firm outcomes is equal to the treatment effect on a business practices index multiplied by the correlation between business practices and this firm outcome. We therefore have:

Prior on Export Outcome = Association between Export Practices and Export Outcome in Pre-Intervention Data * Literature-informed Prior of Treatment Effect on Export Practices.

Panel C of Table 3 in our registered report gives us the association between 2017 export outcomes in our data and our baseline export practices index, along with a standard error. McKenzie and Woodruff (2014) report that business training programs typically increase the proportion of business practices implemented by 0.05 to 0.10. We take a mean of 0.10, since our intervention is more intensive than most business training. This mean is also similar to the 9 percentage point increase in management practices found in Colombia by Iacovone et al. (2019). The standard deviation of the change in export practices is assumed to be 0.03, for a 95% confidence interval of (0.04, 0.16).

Then, for example, to get the literature-informed prior on the effect of exporting at all, we multiply the association in Table 3 (0.907) by the prior on the treatment effect on export practices (0.10) to get a mean for the prior of 0.091. To get the standard error, we assume independence of the errors in the effect on export practices from the error in the association between export practices and export outcomes. Using the formula that Var(AB) = Var(A)Var(B)+Var(A)E(B) +Var(B)E(A) under independence, we then can calculate that the prior standard deviation for the effect on exporting at all is $(0.14^2 0.03^2+0.14^2 0.1+0.03^2 0.907)^{1/2}$

= 0.05. We use similar calculations for the other export outcomes in Table 3 to get the following literature-informed priors:

2. *Impact on whether firms export in the year after the program*: Gaussian with mean of 0.091 and standard deviation of 0.053.

3. *Impact on the number of different products exported in the year after the program starts*: Gaussian with mean of 1.3 and standard deviation of 1.3.

4. *Impact on number of countries exported to in the year after the program starts*: Gaussian with mean of 0.66 and standard deviation of 0.36.

5. *Impact on the Number of Distinct Product-Country Combinations Exported in the Year After the program starts* : Gaussian with mean of 4.5 and standard deviation of 4.4.

6. *Impact on export innovation / the percent of firms that export to a new country-product combination in the year after the program starts*: Gaussian with mean of 8.6 and standard deviation of 5.6.

7. *Impact on export value / growth of dollar value of exports in percentage terms in the year after the program starts*: Inverse hyperbolic sine of exports treatment effect has a prior that is Gaussian with mean of 1.37 and a standard deviation of 0.58.

8. *Impact on export labor productivity of firms in the year after the program starts*: Gaussian with mean of 0.38 and standard deviation of 0.20.

9. *Average impact across different export outcomes measured by the export index in the year after the program starts*: Table 3 did not report the correlation between the export practice index and export practices. The association using the same controls as in panel C of Table 3 is 1.87 with a standard error of 0.26. The prior is therefore a Gaussian with a mean of 0.19 and a standard deviation of 0.09. (Note this is interpretable as the effect in standard deviations, so the mean is for a 0.19 s.d. increase in export practices).

10. *Correlation in export performance*: The literature contains no information about this because empirical data cannot speak directly to this parameter. However, to understand the impact of this correlation on inference about the program, we will perform inference in the case where this correlation is 0, 0.5 and 1.

Default Priors

In the absence of solid evidence in the literature, or in the case of a multi-study aggregation, a Bayesian analysis may lack strong prior information outside the data. However in this case it is still possible to place proper yet weak priors on the parameters of interest using our knowledge of the likely magnitude of these effects. For example, all of the effects we study are more likely to be on the order of 10 or 100 than on the order of 10,000 or 100,000 in the units we measure, and this is evident simply by the scale of variation in the previous years' data (and the logically possible variation given that many outcomes are in percentage terms or standardized incidences).

It is also possible to derive a default location for the priors for treatment effects parameters from the common practice of testing and either “rejecting” or “failing to reject” null hypotheses of zero within a frequentist framework in economics. This roughly corresponds to a system in which effects are assumed to be zero until the data contains sufficient evidence against that assumption. Therefore, effects priors can often be centered at zero to reflect this generally accepted approach.

In certain cases where the parameters have finite support it is possible to put a uniform prior on them corresponding to equal chance of any given parameter value.

This informs the following choice of “default” weakly informative priors for the parameters above:

Question 1 (take-up): A uniform on the interval [0,100].

Questions 2 – 9 (export outcomes): A gaussian with mean 0 and standard deviation 10 times the variation in the previous year’s outcome.

Question 10 (correlation). A uniform on the interval [-1,1].