Supplementary Information

S1 Critique of previous analyses and rationale for new analysis

S1.1 Pseudo-replication

The two most recent analyses of the organic to conventional yield gap utilized the yield ratios without always taking into account the underlying data structure, leading to potential pseudo-replication and an understated Type 1 error rate [1, 2].

The random effect meta-analysis employed by Seufert et al. [1] assumes that effect sizes are independent and drawn from a common distribution [3]. However, multiple response ratios were extracted from studies without nesting these observations. Doing so may violate the assumption of independence, since multiple response ratios from the same study may be non-independent. In addition, treating response ratios from the same study as independent gives the studies with the most response ratios disproportionate weight while also inflating the replication and hence artificially reduces the confidence region of the analysis. Further, in approximately 60% of studies, one set of yield data was used as the baseline for multiple comparisons and these were then treated as independent data points. Treating response ratios that incorporate the same baseline as independent data points also violates assumptions of independence for meta-analysis [4, 5].

The de Ponti et al. [2] study had similar issues with pseudo-replication. Additionally they did not account for the sampling variance within studies, which is the recommended practice to deal with unequal variances in the sample of studies [6].

Because we constructed a random effects model (Eq. S1) and conducted the same categorical comparisons as Seufert et al. [1], we first investigated the effects of not accounting for the hierarchical data structure on the Type 1 error rate for the Seufert et al. model. We also investigated how much of the difference in our results was due to the hierarchical modeling approach, versus the additional data in our study. We did not investigate the de Ponti et al. [2] analysis in detail because its structure was entirely different from ours.

S1.2 Re-analysis of Seufert et al. data

In order to estimate the Type I error rate of the Seufert et al. analysis, we used a randomization test. We forced the null hypothesis to be true by randomly re-assigning the 'organic' and 'conventional' labels for each study and then using the R package `Metafor` [7] to implement a random effects meta-analysis on each randomized dataset. Repeating this procedure 10^5 times enabled us to determine the Type I error rate (false rejection) resulting from not accounting for the hierarchical structure of the data. In over 50% of simulations, the null hypothesis was rejected using a nominal Type I error rate of 0.05 (Fig. S1). In other words, even if organic and conventional yields are known not to be different, applying the model used by [1] for these data would lead to the conclusion that they are significantly different in over 50% of cases. This means that the actual Type I error rate is inflated relative to what was reported, leading to the following related statements: the significance levels were overstated; the confidence intervals were underestimated; the uncertainty was not fully accounted for. This is a likely explanation for why these authors found many significant differences between explanatory variables for management, study quality, and crop type, whereas we did not.

S1.3 How much are differences in results due to the model or the data?

We conducted a series of tests to evaluate the extent to which differences between our results and those of Seufert et al. [1] depended on the model we use or the differences among the datasets (the Seufert et al. data was a subset of the data we used).

First, we re-analyzed Seufert et al.'s [1] data (316 comparisons from 66 studies) with their non-nested model in a Bayesian framework in order to verify that any differences in parameter estimates were not attributable to a change in statistical paradigms. The non-nested, random effects meta-analytic model is:

$$\begin{aligned} y_i &= \mu + \alpha_i + \epsilon_i \\ \alpha_i &\sim N(0, \sigma_\alpha^2) \\ \epsilon_{ijk} &\sim N(0, S_{ijk}) \end{aligned} \tag{S1}$$

where y_i is the observed effect of the i^{th} response ratio, μ is the average true effect, α_i is the random effect of study i, ϵ_i is the residual, σ_α^2 is the between-study variance in true effect sizes [3]. The value of S_{ijk} is estimated by the reported standard error of the effect in study i.

We implemented the non-nested model using a Bayesian framework using JAGS interfaced via the package `rjags` [8, 9]. The yield gap estimate of the Bayesian and frequentist non-nested models using Seufert et al.'s data did not differ quantitatively (compare a and b, Fig. S2).

Next, we re-analyzed Seufert et al.'s [1] data with our nested model, accounting for the hierarchy in the meta-dataset. Our re-analysis shrunk the yield gap by 4 percentage points and increased the uncertainty around these estimates (Fig. S2c).

Next, using our larger data-set we implemented their non-nested model in a Bayesian and frequentist framework. The frequentist analysis was implemented in the R package `metafor`. The non-nested analysis on our more comprehensive dataset gave an estimate of the organic to conventional ratio almost identical to Seufert et al. [1]'s original results using both frequentist (Fig. S2d) and Bayesian approaches (Fig. S2e).

Finally, our model with our expanded dataset (1071 comparisons from 115 studies) shrinks the estimate of the yield gap an additional 2% in comparison to the estimate using Seufert et al.'s data and our model (Fig. S2f). The variance around the estimate also shrinks slightly (compared to c), as expected with a larger data-set. Thus, difference in methods is twice as important as the difference in data in determining the reduction in the estimate of the organic to conventional yield gap between Seufert et al.'s analysis and ours.

S2 Building our meta-analytic model

S2.1 Modeling framework

We chose to use a Bayesian modeling framework because the existing Markov chain Monte Carlo sampling methods [8] enabled us handle the complexity of the hierarchical model we constructed. Such methods did not exist for frequentist analyses at the time the analyses were conducted. Because the posterior is the product of the likelihood and the prior, when using uninformative (flat) priors as we have done here, the posterior should be approximately calibrated to frequentist results.

S2.2 Effect size

We chose to use the response ratio (the ratio of the mean outcome of the treatments of interest) because it quantifies the proportional difference between the treatments (in this case, the organic and conventional yields). In our analyses we used the natural logarithm of the response ratio because (1) it has an approximately normal sampling distribution, whereas the sampling distribution of the raw response ratio is skewed, and (2) deviations in the numerator and the denominator hold equal weight [10]. We then back transformed the model output to facilitate interpretation.

S2.3 Random/mixed model

In a random and mixed effects meta-analytic models, the true study effects are assumed to come from a common distribution and, thus, such models provide inferences about the larger population of possible studies from which those included in the actual analysis are a random sample [11].

In contrast to random and mixed effects models, fixed effect models assumes there is little heterogeneity of effect size estimates between studies. The study effects are therefore not modeled as being drawn from a common distribution. We did not conduct any fixed effect analyses because, when using fixed-effects models, the goal is to make a conditional inference only about the studies included in the meta-analysis [11].

S2.4 Parameter inclusion

To determine the levels of hierarchy supported by the data, we sequentially added random effects and examined the posteriors of the parameters to determine the support for their inclusion. We also confirmed our selection with Deviance Information Criterion (DIC). The DIC can be problematic for hierarchical models because the effective number of parameters is not clearly defined [12, 13]. The DIC was therefore used in combination with a visual examination of the posterior distributions of the parameters to select the best supported model.

For the variance within and between year random effect distributions, we considered two parameterizations: 1) the variance terms, denoted σ_η^2 and σ_β^2, respectively, were shared across all studies each with a Uniform(0,100) prior, and 2) the variance terms were study-specific (i.e., $\sigma_\eta^2[i]$ and $\sigma_\beta^2[i]$). In the latter case, the study-specific precision terms (1/variance) were assumed to be distributed according to a gamma distribution whose parameters were estimated. Uniform(0,100) priors were used for the coefficient of variation $(1/\sqrt{shape})$ and the square root of the scale.

We first added a random effect of study and examined the posterior for σ_α (the standard deviation of the common distribution from which the study effects are drawn). The posterior was clearly differentiated from zero (Fig. S6a). We next added random variation within a year and examined σ_η. We found it was also clearly different from zero (Fig. S6b). The DIC was also smaller than when only a random effect of study was included (Tab. S1). Next we allowed the within-year precisions to be study-specific and follow a gamma distribution. We examined the coefficient of variation of the gamma distribution and found it was clearly differentiated from zero (Fig. S6c). The DIC was also smaller than when a single within year effect was shared across studies (Tab. S1). Lastly, we added a between year random effect and examined σ_β. The posterior was concentrated at zero (Fig. S6d) so we concluded there was insufficient support for including it in the model. The estimate of the yield gap and its uncertainty did not differ substantially from when no between year effect was included (Fig. S7), the DIC, however, was marginally smaller then when no between year random effect was included (Tab. S1).

S2.5 Weighting

The estimates of effect size from different studies will differ in their precision, or standard error. To handle these differences in precision, the sampling variance from each study are used as an estimate of the precision of the response ratio [3, 6, 10, 11, 14]. The variance of a response ratio is equal to

$$\frac{SD_{org}^2}{N_{org}\overline{X}_{org}^2} + \frac{SD_{conv}^2}{N_{conv}\overline{X}_{conv}^2} \tag{S2}$$

Where SD is the standard deviation, $\overline{X}$ is the mean and N is the sample size of the organic and conventional treatments [10]. This weighting increases the power of the test and the precision of the combined estimate [6]. We therefore use the estimated variance of log response ratios [10] to weight our effect sizes.

We tested the sensitivity of the analysis to weighting by implementing a non-nested, unweighted model using a frequentist framework [7] (our nested model does accommodate unweighted variances). The mean estimate of the unweighted, non-nested model was the same as the weighted, non-nested model, but the confidence intervals around the parameter estimate were slightly smaller (Fig. S2g).

In many multi-year studies, we could not obtain the within-year variance among observations. In these cases, the variance of the means across years was used as an estimate of the observation-level sampling variance (which was also how [1] combined such data). Because this variance estimate lumps the between-year variation and the sampling variation together, it is an overestimate of the sampling variance. For example, for a single study with multiple years,

$$y_i = \mu + \beta_i + \epsilon_i, i = 1, \ldots, n \tag{S3}$$

$$\epsilon_i \sim N(0, \sigma^2_{\epsilon,i}) \tag{S4}$$

$$\beta_i \sim N(0, \sigma^2_\beta) \tag{S5}$$

where y_i is the log response ratio of the i^{th} year of a study, μ is the average, ϵ_i is the residual due to sampling variation, $\sigma^2_{\epsilon,i}$ is the sampling variance in year i, β_i is the random effect of the i^{th} year, and σ^2_β is the variance of the distribution from which the year effects are drawn. The sample variance of the y_i is an estimate of the between-year variance (σ^2_β) plus the average of the sampling variances ($\sigma^2_{\epsilon,i}$). In a mixed or random effects meta-analysis, this becomes the estimate of the within-study variance (σ^2_i).

We conducted the hierarchical meta-analysis with and without the studies that reported only a between-year variance to examine the potential bias that under-weighting some studies might introduce. Excluding the studies that only provided the between-year variance decreased our meta-dataset to 56 studies and 839 observations. The yield gap estimated from the subset of the data was only 2 percentage point larger than the estimate from the full dataset with a 1.6 percentage point increase in the estimate of the variance (S2, compare f and h), as expected with less data. Including the studies without a true observation-level sampling variance estimate, therefore, does not substantially change our estimate of the yield gap.

S2.6 Sampling dependence

Studies included in meta analyses often employ ANOVA-style designs where multiple treatments are compared against a single control. In these cases, a response ratio can be calculated for each control-treatment pair. These response ratios, however, are not independent, because they share a common control and therefore should not be included separately in an analysis [4, 5]. When response ratios shared a common control, we calculated a combined response ratio (y_{ijk}) and corresponding standard error (σ_{ijk}) for the entire study using the method presented in [Eq. 3 & 8, 4]. In the analysis with no explanatory variables, 63% of the data were combined using the Lajeunesse method.

If the response ratios are grouped separately among categories in a mixed effects model the bias introduced by non-independence is minimized because the response ratios are not pooled together [4]. Therefore we did not aggregate response ratios in the analyses including explanatory variables when the response ratios were in different levels of the explanatory variable.

S3 Meta-datasets and publication bias

The inclusion criteria and the timing of literature searches differed between our study and that of de Ponti et al. [2] and Seufert et al. [1] and thus the meta-data sets vary in their coverage (Fig. S8). We used similar inclusion criteria to Seufert et al. but excluded (1) comparisons of organic yields with subsistence yields because the latter do not represent conventional agriculture, and (2) comparisons of yields not from the same year. Our meta-dataset encompasses 115 studies published between 1977 and 2012, of which Seufert et al.'s meta-data set is a subset (66 studies from 1980 to 2010). Unlike our study and that of

Seufert et al., de Ponti et al. included studies that did not report an estimate of sampling variance. de Ponti's less strict exclusion criteria enabled them to include 156 studies from 1989 to 2010.

For all three studies, we found that study publication year is correlated with the yield ratio (Fig. S9). Specifically, there is a trend towards larger yield gaps with conventional out-performing organic in more recent studies. This effect is most pronounced in the data used by Seufert et al., but is also present in de Ponti et al.'s and our own meta-dataset. This trend could result from conventional yields increasing relative to organic yields through time. Historically, research and development of organic cropping systems has been extensively underfunded relative to conventional systems [15–17], so it is not inconceivable that continued investment in conventional techniques has widened the yield gap through time.

The same trend could also result if a publication bias favoring studies that report higher conventional yields has increased through time. Not all studies are submitted for publication and, of those that are, not all are accepted. Publication bias will result if studies that show significant results are preferentially submitted and published, or if studies are suppressed because the findings do not align with the interests of the researchers or funding sources [18]. A bias would also result if researchers are choosing study systems which they expect to show larger differences, and thus there is a bias in what studies are conducted. Conversely, a bias would occur if the crop species of greatest interest to researchers is one that exhibits a large yield gap (e.g., cereals). Interesting, the proportion of studies on cereals in the literature has increased through time (Fig. S10), which could explain both the bias toward studies that report higher conventional yields and the increase in the gap through time.

We visually assessed bias in our dataset and in the subset of data comprising the Seufert et al. data using a funnel plot and QQ-plot [6]. Asymmetrical funnel plots may indicate a systematic difference between smaller and larger studies ("small study effects") which may be due to publication bias [6, 18]. Our funnel plot revealed a slight asymmetry favoring small studies that report conventional yields are higher relative to organic, Fig. S11a). The QQ-plot confirmed this observation: the observed quantiles are first below then above the expected line, suggesting the observed data is gaining quantiles faster than expected under a Gaussian model (Fig. S11b). This is because in the observed quantiles, there is more mass in the tail compared to normally distributed data (i.e., the distribution of response ratios is fat-tailed). That the lower quantiles are further from the line and even steeper suggests that there is more mass in the lower tail (i.e., where conventional yields are higher than organic). A fat-tailed distribution of response ratios could be due to bias. Funnel and QQ-plots of the subset of data used by Seufert et al. were similar.

All of these assessments suggest that there is bias in the literature favoring studies that report conventional yields out-performing organic. Our results should therefore be viewed as a potential overestimate of the yield gap. It is unclear, however, whether this bias is due to publication bias or another type of bias such as a bias in what studies were conducted.

S4 Explanatory variable inclusion

To examine the support for including different explanatory variables in a model, various forms of model selection can be employed to compare models that contain different combinations of explanatory variables (e.g., [19]). Because the studies in our analysis did not consistently report many characteristics we wished to include in our model (e.g., nitrogen input, rotation type, tillage), we were unable to use such an approach, as it requires that the data are consistent across models. We instead analyzed each explanatory variable separately. Although not ideal, it is not possible to include all explanatory variables investigated here in one model and still use a traditional model selection framework. Organic and conventional yields were deemed significantly different from each other if the 95% credible interval of the yield ratio did not overlap one. Different levels of explanatory variables were considered to be significantly different if the posterior of the 95% credible interval of the difference between the group means did not overlap zero.

References

[1] Seufert, V., Ramankutty, N. & Foley, J. A. 2012 Comparing the yields of organic and conventional agriculture. *Nature*, **485**(7397), 229–232.

[2] de Ponti, T., Rijk, B. & van Ittersum, M. 2012 The crop yield gap between organic and conventional agriculture. *Agricultural Systems*, **108**, 1–9.

[3] Hedges, L. V. & Olkin, I. 1985 *Statistical methods for meta-analysis*. New York: Academic Press.

[4] Lajeunesse, M. J. 2011 On the meta-analysis of response ratios for studies with correlated and multi-group designs. *Ecology*, **92**(11), 2049–2055.

[5] Gleser, L. J. & Olkin, I. 1994 Stochastically dependent effect sizes. In *The handbook of research synthesis* (eds. H. Cooper & L. V. Hedges), pp. 339–355. New York: Russell Sage Foundation.

[6] Gurevitch, J. & Hedges, L. V. 1999 Statistical issues in ecological meta-analyses. *Ecology*, **80**(4), 1142–1149.

[7] Viechtbauer, W. 2010 Conducting meta-analyses in r with the metafor package. *Journal of Statistical Software*, **36**(3), 1–48.

[8] Plummer, M. 2013 *rjags: Bayesian graphical models using MCMC*. R package version 3–10.

[9] Plummer, M. 2003 Jags: A program for analysis of bayesian graphical models using gibbs sampling.

[10] Hedges, L., Gurevitch, J. & Curtis, P. 1999 The meta-analysis of response ratios in experimental ecology. *Ecology*, **80**, 1150–1156.

[11] Hedges, L. & Vevea, J. 1998 Fixed- and random-effects models in meta-analysis. *Psychological Methods*, **3**(4), 486–504.

[12] Gelman, A. & Hill, J. 2006 *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press.

[13] Kéry, M. & Schaub, M. 2012 *Bayesian population analysis using WinBUGS: a hierarchical perspective*. Academic Press.

[14] Gurevitch, J. & Hedges, L. 2001 Meta-analysis. combining the results of independent experiments. In *Design and analysis of ecological experiments* (eds. S. M. Scheiner & J. Gurevitch), pp. 347–369. New York: Oxford University Press.

[15] Vanloqueren, G. & Baret, P. V. 2009 How agricultural research systems shape a technological regime that develops genetic engineering but locks out agroecological innovations. *Research policy*, **38**(6), 971–983.

[16] McIntyre, B. D. 2009 *IAASTD International Assessment of Agricultural Knowledge, Science and Technology for Development: Global Report*. Island Press.

[17] Carlisle, L. & Miles, A. 2013 Closing the knowledge gap: How the usda could tap the potential of biologically diversified farming systems. *Journal of Agriculture, Food Systems, and Community Development*, **3**(4), 219–225.

[18] Rothstein, H. R., Sutton, A. J. & Borenstein, M. 2006 *Publication bias in meta-analysis: Prevention, assessment and adjustments*. John Wiley & Sons.

[19] Johnson, J. B. & Omland, K. S. 2004 Model selection in ecology and evolution. *Trends in Ecology & Evolution*, **19**(2), 101–108.

[20] Badgley, C., Moghtader, J., Quintero, E., Zakem, E., Chappell, M. J., Aviles-Vazquez, K., Samulon, A. & Perfecto, I. 2007 Organic agriculture and the global food supply. *Renewable agriculture and food systems*, **22**(2), 86–108.

[21] Amarante, C. V. T. d., Steffens, C. A., Mafra, Á. L. & Albuquerque, J. A. 2008 Yield and fruit quality of apple from conventional and organic production systems. *Pesquisa Agropecuária Brasileira*, **43**(3), 333–340.

[22] Appireddy, G. K., Saha, S., Mina, B. L., Kundu, S., Selvakumar, G. & Gupta, H. S. 2008 Effect of organic manures and integrated nutrient management on yield potential of bell pepper (*Capsicum annuum*) varieties and on soil properties. *Archives of Agronomy and Soil Science*, **54**(2), 127–137.

[23] Archer, D. W., Jaradat, A. A., Johnson, J. M., Weyers, S. L., Gesch, R. W., Forcella, F. & Kludze, H. K. 2007 Crop productivity and economics during the transition to alternative cropping systems. *Agronomy journal*, **99**(6), 1538–1547.

[24] Arncken, C. M., Mäder, P., Mayer, J. & Weibel, F. P. 2012 Sensory, yield and quality differences between organically and conventionally grown winter wheat. *Journal of the Science of Food and Agriculture*, **92**(14), 2819–2825.

[25] Ban, D., Oplanić, M., Ilak Peršurić, A. S., Novak, B., Žutić, I., BoroŠić, J. & ŽnidarČiČ, D. 2007 Crop management systems and endomycorrhiza effects on endive (*Cichorium endivia L.*) growth. *Acta agriculturae Slovenica*, **89**(1), 35–43.

[26] Behera, B., Sankar, G. M., Sharma, K., Mishra, A., Mohanty, S., Mishra, P., Rath, B. & Grace, J. K. 2012 Effects of fertilizers on yield, sustainability, and soil fertility under rainfed pigeon pea + rice system in subhumid oxisol soils. *Communications in Soil Science and Plant Analysis*, **43**(17), 2228–2246.

[27] Bertschinger, L., Mouron, P., Dolega, E., Höhn, H., Holliger, E., Husistein, A., Schmid, A., Siegfried, W., Widmer, A. *et al.* 2004 Ecological apple production: a comparison of organic and integrated apple-growing. In *Acta Horticulturae 638, XXVI International Horticultural Congress: Sustainability of Horticultural Systems in the 21st Century, Toronto, Canada*, pp. 321–332. International Society for Horticultural Science.

[28] Besson, J., Oberson, A., Michel, V. & Niggli, U. 1992 Dok-versuch: vergleichende langzeituntersuchungen in den drei anbausystemen biologisch-dynamisch, organisch-biologisch und konventionell. ii. ertrag der kulturen: Gerste, 1. und 2. fruchtfolgeperiode. *Schweiz. Landw. Fo.*, **32**, 3–32.

[29] Besson, J., Michel, V., Spiess, E. & Niggli, U. 1993 Dok-versuch: vergleichende langzeit-untersuchungen in den drei anbausystemen biologisch-dynamisch, organisch-biologisch und konventionell. iv. aufwand und ertrag: Naehrstoffbilanzen, 1. und 2. fruchtfolgeperiode. *Schweiz. Landw. Fo*, **32**, 199–218.

[30] Besson, J., Meyre, S., Spiess, E., Stauffer, W. & Niggli, U. 1993 Dok-versuch: vergleichende langzeituntersuchungen in den drei anbausystemen biologisch- dynamisch, organisch-biologisch und konventionell. ii. ertrag der kulturen: Randen, 2. fruchtfolgeperiode. *Schweiz. Landw. Fo*, **32**, 449–463.

[31] Bettiol, W., Ghini, R., Galvão, J. A. H. & Siloto, R. C. 2004 Organic and conventional tomato cropping systems. *Scientia Agricola*, **61**(3), 253–259.

[32] Bicanová, E., Capouchová, I., Krejícová, L., Petr, J. & Erhartová, D. 2006 The effect of growth structure on organic winter wheat quality. *Zemdirbyste. Mokslo Darbai*, **93**, 297–305.

[33] Blaise, D. 2006 Yield, boll distribution and fibre quality of hybrid cotton (*Gossypium hirsutum L.*) as influenced by organic and modern methods of cultivation. *Journal of Agronomy and Crop Science*, **192**(4), 248–256.

[34] Campanelli, G. & Canali, S. 2012 Crop production and environmental effects in conventional and organic vegetable farming systems: The case of a long-term experiment in mediterranean conditions (Central Italy). *Journal of Sustainable Agriculture*, **36**(6), 599–619.

[35] Cavigelli, M. A., Hima, B. L., Hanson, J. C., Teasdale, J. R., Conklin, A. E. & Lu, Y.-c. 2009 Long-term economic performance of organic and conventional field crops in the mid-Atlantic region. *Renewable agriculture and food systems*, **24**(2), 102–119.

[36] Citak, S. & Sonmez, S. 2010 Effects of conventional and organic fertilization on spinach (*Spinacea oleracea, L.*) growth, yield, vitamin C and nitrate concentration during two successive seasons. *Scientia Horticulturae*, **126**(4), 415–420.

[37] Clark, S., Klonsky, K., Livingston, P. & Temple, S. 1999 Crop-yield and economic comparisons of organic, low-input, and conventional farming systems in California's Sacramento Valley. *American Journal of Alternative Agriculture*, **14**(3), 109–121.

[38] Coulter, J. A., Sheaffer, C. C., Wyse, D. L., Haar, M. J., Porter, P. M., Quiring, S. R. & Klossner, L. D. 2011 Agronomic performance of cropping systems with contrasting crop rotations and external inputs. *Agronomy Journal*, **103**(1), 182–192.

[39] Cürük, S., Sermenli, T., Mavi, K. & Evrendilek, F. 2004 Yield and fruit quality of watermelon (*Citrullus lanatus* (thumb.) *Matsum*. & *Nakai*.) and melon (*Cucumis melo, L.*) under protected organic and conventional farming systems in a Mediterranean region of Turkey. *Biological agriculture & horticulture*, **22**(2), 173–183.

[40] De Luca, S., Fagnano, M. & Quaglietta Chiarandà, F. 2004 The effect of organic fertilization on yields of tomato crops in the Sele River Plain. In *International Symposium Towards Ecologically Sound Fertilisation Strategies for Field Vegetable Production* (eds. F. Tei, P. Benincasa & M. Guiducci), vol. 700, pp. 103–106.

[41] Delate, K. & Cambardella, C. A. 2004 Agroecosystem performance during transition to certified organic grain production. *Agronomy Journal*, **96**(5), 1288–1298.

[42] Delate, K., Duffy, M., Chase, C., Holste, A., Friedrich, H. & Wantate, N. 2003 An economic comparison of organic and conventional grain crops in a long-term agroecological research (LTAR) site in Iowa. *American Journal of Alternative Agriculture*, **18**(2), 59–69.

[43] Delmotte, S., Tittonell, P., Mouret, J.-C., Hammond, R. & Lopez-Ridaura, S. 2011 On farm assessment of rice yield variability and productivity gaps between organic and conventional cropping systems under mediterranean climate. *European Journal of Agronomy*, **35**(4), 223–236.

[44] Demiryurek, K. & Ceyhan, V. 2009 Economics of organic and conventional hazelnut production in the Terme district of Samsun, Turkey. In *VII International Congress on Hazelnut, Viterbo, Italy, 23-27 June 2008* (eds. L. Varvaro & S. Franco), 845, pp. 739–744. International Society for Horticultural Science (ISHS).

[45] Denison, R. F., Bryant, D. C. & Kearney, T. E. 2004 Crop yields over the first nine years of LTRAS, a long-term comparison of field crop systems in a Mediterranean climate. *Field Crops Research*, **86**(2), 267–277.

[46] Dobbs, T. L. & Smolik, J. D. 1997 Productivity and profitability of conventional and alternative farming systems: A long-term on-farm paired comparison. *Journal of Sustainable Agriculture*, **9**(1), 63–79.

[47] Doltra, J., Lægdsmand, M. & Olesen, J. E. 2011 Cereal yield and quality as affected by nitrogen availability in organic and conventional arable crop rotations: A combined modeling and experimental approach. *European Journal of Agronomy*, **34**(2), 83–95.

[48] Drinkwater, L., Letourneau, D., Workneh, F., Van Bruggen, A. & Shennan, C. 1995 Fundamental differences between conventional and organic tomato agroecosystems in California. *Ecological Applications*, pp. 1098–1112.

[49] Drinkwater, L., Janke, R. & Rossoni-Longnecker, L. 2000 Effects of tillage intensity on nitrogen dynamics and productivity in legume-based grain systems. *Plant and Soil*, **227**(1-2), 99–113.

[50] Eckhoff, J., Bergman, J. & Flynn, C. 2005 A comparison of safflower (*Carthamus tinctorius L.*) grown under conventional and chemical-free conditions. In *Proceedings of the VIth International Safflower Conference, Istanbul-Turkey, 6-10 June, 2005. SAFFLOWER: a unique crop for oil spices and health consequently, a better life for you.* (eds. E. Esendal, J. Bergman, N. Kandemir, R. Johnson & A. Corleto), pp. 3–1. Engin Maatbacilik Ltd. Şti.

[51] Efthimiadou, A., Bilalis, D., Karkanis, A., Froud-Williams, B. & Eleftherochorinos, I. 2009 Effects of cultural system (organic and conventional) on growth, photosynthesis and yield components of sweet corn (*Zea mays L.*) under semi-arid environment. *Notulae Botanicae Horti Agrobotanici Cluj-Napoca*, **37**(2), 104–111.

[52] Entz, M., Hoeppner, J., Wilson, L., Tenuta, M., Bamford, K. & Holliday, N. 2005 Influence of organic management with different crop rotations on selected productivity parameters in a long-term Canadian field study. In *Researching Sustainable Systems, Proceedings of the International Scientific Conference on Organic Agriculture*.

[53] Eyhorn, F., Ramakrishnan, M. & Mäder, P. 2007 The viability of cotton-based organic farming systems in India. *International Journal of Agricultural Sustainability*, **5**(1), 25–38.

[54] Gelfand, I., Snapp, S. S. & Robertson, G. P. 2010 Energy efficiency of conventional, organic, and alternative cropping systems for food and fuel at a site in the US Midwest. *Environmental science & technology*, **44**(10), 4006–4011.

[55] Gerdgikova, M., Videva, M. & Pavlov, D. 2012 Content and yield of crude protein from winter pea grain, cultivated after different predecessors in conditions of organic and conventional production. *Agricultural Science and Technology*, **4**(4), 378–381.

[56] Gliessman, S., Werner, M., Swezey, S., Caswell, E., Cochran, J. & Rosado-May, F. 1996 Conversion to organic strawberry management changes ecological processes. *California Agriculture*, **50**(1), 24–31.

[57] Goldstein, W., Barber, W., Carpenter-Boggs, L., Daloren, D. & Koopmans, C. 2004 Comparisons of conventional, organic and biodynamic methods. *Michael Fields Agricultural Institute*.

[58] Gopinath, K., Saha, S., Mina, B., Pande, H., Kundu, S. & Gupta, H. 2008 Influence of organic amendments on growth, yield and quality of wheat and on soil properties during transition to organic production. *Nutrient cycling in agroecosystems*, **82**(1), 51–60.

[59] Gosavi, A., Potdar, D., Sonawane, P., Shirpurkar, G. & Rasal, P. 2009 Organic farming in soybean-wheat cropping sequence. *Agricultural Science Digest*, **29**(4), 267–270.

[60] Greer, G., Kaye-Blake, W., Zellman, E. & Parsonson-Ensor, C. 2008 Comparison of

the financial performance of organic and conventional farms. *Journal of Organic Systems*, **3**(2), 18–28.

[61] Gündoğmuş, E. 2006 A comparative analysis of organic and conventional dried apricot production on small households in Turkey. *Asian Journal of Plant Sciences*, **5**(1), 98–104.

[62] Gündoğmuş, E. 2007 Organic dried fig production: A comparative analysis of organic and conventional smallholdings in Turkey. *Biological Agriculture & Horticulture*, **24**(4), 379–396.

[63] Gündoğmuş, E. & Bayramoglu, Z. 2005 Organic raisin production: a comparative analysis of organic and conventional smallholdings in Turkey. *Journal of Agronomy*, **4**.

[64] Hargreaves, J. C., Adl, M., Warman, P. R. & Rupasinghe, H. 2008 The effects of organic and conventional nutrient amendments on strawberry cultivation: Fruit yield and quality. *Journal of the Science of Food and Agriculture*, **88**(15), 2669–2675.

[65] Helmers, G. A., Langemeier, M. R. & Atwood, J. 1986 An economic analysis of alternative cropping systems for east-central Nebraska. *American Journal of Alternative Agriculture*, **1**(4), 153–158.

[66] Herencia, J. F., Ruiz-Porras, J., Melero, S., Garcia-Galavis, P., Morillo, E. & Maqueda, C. 2007 Comparison between organic and mineral fertilization for soil fertility levels, crop macronutrient concentrations, and yield. *Agronomy Journal*, **99**(4), 973–983.

[67] Institute for Biodynamic Research (IBDF) 1988-1991 growth and yield of winter rye.

[68] Igbokwe, P. E., Huam, L. C., Chukwuma, F. O. & Huam, J. 2006 Sweetpotato yield

and quality as influenced by cropping systems. *Journal of vegetable science*, **11**(4), 35–46.

[69] Ingver, A., Tamm, I. & Tamm, Ü. 2008 Effect of organic and conventional production on yield and the quality of spring cereals. *Agronomijas Vēstis*, **11**, 61–67.

[70] Järvan, M. & Edesi, L. 2009 The effect of cultivation methods on the yield and biological quality of potato. *Agronomy Research*, **7**, 289–299.

[71] Jimenez, M., Van der Veken, L., Neirynck, H., Rodríguez, H., Ruiz, O. & Swennen, R. 2007 Organic banana production in Ecuador: Its implications on black Sigatoka development and plant–soil nutritional status. *Renewable Agriculture and Food Systems*, **22**(04), 297–306.

[72] Juroszek, P., Ledesma, D., Ma, C., Yang, R., Lumpkin, H., Lin, C., Tsai, H., Wu, D., Hanson, P. *et al.* 2008 Plant vigour and yields of organically and conventionally grown tomato crops in Taiwan. *Acta horticulturae*, **767**, 257–265.

[73] Kalinova, J. & Vrchotova, N. 2011 The influence of organic and conventional crop management, variety and year on the yield and flavonoid level in common buckwheat groats. *Food chemistry*, **127**(2), 602–608.

[74] Kaut, A., Mason, H., Navabi, A., O'donovan, J. & Spaner, D. 2009 Performance and stability of performance of spring wheat variety mixtures in organic and conventional management systems in western Canada. *The Journal of Agricultural Science*, **147**(02), 141–153.

[75] Kirchmann, H., Bergström, L., Kätterer, T., Mattsson, L. & Gesslein, S. 2007 Comparison of long-term organic and conventional crop–livestock systems on a previously nutrient-depleted soil in Sweden. *Agronomy Journal*, **99**(4), 960–972.

[76] Kitchen, J. L., McDonald, G. K., Shepherd, K. W., Lorimer, M. F. & Graham, R. D. 2003 Comparing wheat grown in South Australian organic and conventional farming systems. 1. growth and grain yield. *Crop and Pasture Science*, **54**(9), 889–901.

[77] Klepper, R., Lockeretz, W., Commoner, B., Gertler, M., Fast, S., O'Leary, D. & Blobaum, R. 1977 Economic performance and energy intensiveness on organic and conventional farms in the Corn Belt: a preliminary comparison. *American Journal of Agricultural Economics*, **59**(1), 1–12.

[78] Knudsen, M. T., Yu-Hui, Q., Yan, L. & Halberg, N. 2010 Environmental assessment of organic soybean (*Glycine max.*) imported from China to Denmark: a case study. *Journal of Cleaner Production*, **18**(14), 1431–1439.

[79] Koocheki, A., Jahan, M. & Nassiri Mahallati, M. 2008 Effects of arbuscular mycorrhizal fungi and free-living nitrogen-fixing bacteria on growth characteristics of corn (*Zea mays L.*) under organic and conventional cropping systems.

[80] Korsaeth, A. 2012 N, P, and K budgets and changes in selected topsoil nutrients over 10 years in a long-term experiment with conventional and organic crop rotations. *Applied and Environmental Soil Science*, **2012**.

[81] Liebhardt, W., Andrews, R., Culik, M., Harwood, R., Janke, R., Radke, J. & Reiger-Schwartz, S. 1989 Crop production during conversion from conventional to low-input methods. *Agronomy Journal*, **81**(2), 150–159.

[82] Lo Scalzo, R., Iannoccari, T., Genna, A., Di Cesare, L., Viscardi, D., Ferrari, V. & Campanelli, G. 2008 Organic vs. conventional field trials: the effect on cauliflower quality. In *Cultivating the Future Based on Science: 2nd Conference of the International Society of Organic Agriculture Research ISOFAR, Modena, Italy, June 18-20, 2008.*

[83] Lockeretz, W., Shearer, G., Sweeney, S., Kuepper, G., Wanner, D. & Kohl, D. H. 1980 Maize yields and soil nutrient levels with and without pesticides and standard commercial fertilizers. *Agronomy Journal*, **72**(1), 65–72.

[84] Lotter, D., Seidel, R. & Liebhardt, W. 2003 The performance of organic and conventional cropping systems in an extreme climate year. *American Journal of Alternative Agriculture*, **18**(3), 146–154.

[85] Long Term Research on Agricultural Systems (LTRAS). Russell Ranch Sustainable Agriculture Facility.

[86] Lyngbaek, A., Muschler, R. & Sinclair, F. 2001 Productivity and profitability of multistrata organic versus conventional coffee farms in Costa Rica. *Agroforestry Systems*, **53**(2), 205–213.

[87] Macit, I., Koc, A., Guler, S. & Deligoz, I. 2007 Yield, quality and nutritional status of organically and conventionally-grown strawberry cultivars. *Asian Journal of Plant Sciences*, **6**(7), 1131–1136.

[88] Mäder, P., Fliessbach, A., Dubois, D., Gunst, L., Fried, P. & Niggli, U. 2002 Soil fertility and biodiversity in organic farming. *Science*, **296**(5573), 1694–1697.

[89] Maggio, A., Carillo, P., Bulmetti, G. S., Fuggi, A., Barbieri, G. & De Pascale, S. 2008 Potato yield and metabolic profiling under conventional and organic farming. *European Journal of Agronomy*, **28**(3), 343–350.

[90] Mahoney, P. R., Olson, K. D., Porter, P. M., Huggins, D. R., Perillo, C. A. & Crookston, R. K. 2007 Profitability of organic cropping systems in southwestern Minnesota. In *Organic Food*, pp. 65–81. Springer.

[91] Malusà, E., Laurenti, E., Ghibaudi, E. & Rolle, L. 2002 Influence of organic and conventional management on yield and composition of grape cv.'Grignolino'. In *XXVI International Horticultural Congress: Viticulture-Living with Limitations 640*, pp. 135–141.

[92] Martínez-Sánchez, J. C. 2008 The role of organic production in biodiversity conservation in shade coffee plantations. Ph.D. thesis, University of Washington.

[93] Martini, E. A., Buyer, J. S., Bryant, D. C., Hartz, T. K. & Denison, R. F. 2004 Yield increases during the organic transition: improving soil quality or increasing experience? *Field Crops Research*, **86**(2), 255–266.

[94] Mazzoncini, M., Barberi, P., Belloni, P., Cerrai, D. & Antichi, D. 2006 Sunflower under conventional and organic farming systems: results from a long term experiment in Central Italy. *Aspects of Applied Biology 79, What will organic farming deliver? COR 2006*, pp. 125–129.

[95] Mazzoncini, M., Belloni, P., Risaliti, R. & Antichi, D. 2007 Organic vs conventional winter wheat quality and organoleptic bread test. In *Improving Sustainability in Organic and Low Input Food Production Systems, University of Hohenheim, Germany*.

[96] Mourão, I., Brito, L. M. & Coutinho, J. 2008 Yield and quality of organic versus conventional potato crop.

[97] Oplanić, M., Ban, D., Ilak Peršurić, A. S. & Žnidarčič, D. 2009 Profitability of leek (*Allium porrum L.*) in three production systems. *International journal of food, agriculture and environment*, **7**(3-4), 376–381.

[98] Peck, G. M., Andrews, P. K., Reganold, J. P. & Fellman, J. K. 2006 Apple orchard productivity and fruit quality under organic, conventional, and integrated management. *HortScience*, **41**(1), 99–107.

[99] Peck, G. M., Merwin, I. A., Brown, M. G. & Agnello, A. M. 2010 Integrated and organic fruit production systems for Liberty apple in the Northeast United States: a systems-based evaluation. *HortScience*, **45**(7), 1038–1048.

[100] Pezzarossa, B., Barbafieri, M., Benetti, A., Petruzzelli, G., Mazzoncini, M., Bonari, E. & Pagliai, M. 1995 Effects of conventional and alternative management systems on soil phosphorus content, soil structure, and corn yield. *Communications in Soil Science & Plant Analysis*, **26**(17-18), 2869–2885.

[101] Pieper, J. R. & Barrett, D. M. 2009 Effects of organic and conventional production systems on quality and nutritional parameters of processing tomatoes. *Journal of the Science of Food and Agriculture*, **89**(2), 177–194.

[102] Polat, E., Demir, H. & Onus, A. 2008 Comparison of some yield and quality criteria in organically and conventionally-grown lettuce. *African Journal of Biotechnology*, **7**(9), 1235–1239.

[103] Polat, E., Demir, H. & Erler, F. 2010 Yield and quality criteria in organically and conventionally grown tomatoes in Turkey. *Scientia Agrícola*, **67**(4), 424–429.

[104] Porter, P. M., Huggins, D. R., Perillo, C. A., Quiring, S. R. & Crookston, R. K. 2003 Organic and other management strategies with two-and four-year crop rotations in Minnesota. *Agronomy Journal*, **95**(2), 233–244.

[105] Posner, J., Baldock, J. & Hedtcke, J. 2005 The Wisconsin integrated cropping systems trials: yields, yield variability, and yield trends 1990-2002. *The Wisconsin Integrated Cropping Systems Trial Tenth Report-2003 & 2004*, p. 66.

[106] Rader, J. S., Walser, R. H., Williams, C. F. & Davis, T. D. 1985 Organic and conventional peach production and economics. *Biological Agriculture & Horticulture*, **2**(3), 215–222.

[107] Raupp, D. 1996 Quality of plant products grown with manure fertilization. pp. 13–33. Institute for Biodynamic Research Darmstadt, Germany.

[108] Raupp, J. 1999 Entwicklung des kornertrages und der ertragskomponenten von winterroggen in einem langzeit-düngungsversuch. *Merbach, W*, pp. 237–240.

[109] Reganold, J. P., Elliott, L. F. & Unger, Y. L. 1987 Long-term effects of organic and conventional farming on soil erosion. *Nature*, **330**(6146), 370–372.

[110] Reganold, J. P., Glover, J. D., Andrews, P. K. & Hinman, H. R. 2001 Sustainability of three apple production systems. *Nature*, **410**(6831), 926–930.

[111] Riahi, A., Hdider, C., Sanaa, M., Tarchoun, N., Ben Kheder, M. & Guezal, I. 2009 Effect of conventional and organic production systems on the yield and quality of field tomato cultivars grown in tunisia. *Journal of the Science of Food and Agriculture*, **89**(13), 2275–2282.

[112] Russo, V. & Taylor, M. 2006 Soil amendments in transition to organic vegetable production with comparison to conventional methods: Yields and economics. *HortScience*, **41**(7), 1576–1583.

[113] Ryan, M. H., Derrick, J. & Dann, P. 2004 Grain mineral concentrations and yield of wheat grown under organic and conventional management. *Journal of the Science of Food and Agriculture*, **84**(3), 207–216.

[114] Ryan, M., Smith, R., Mortensen, D., Teasdale, J., Curran, W., Seidel, R. & Shumway, D. 2009 Weed–crop competition relationships differ between organic and conventional cropping systems. *Weed Research*, **49**(6), 572–580.

[115] Sellen, D., Tolman, J. H., McLeod, D. G. R., Weersink, A. & Yiridoe, E. K. 1996 A

comparison of financial returns during early transition from conventional to organic vegetable production. *Journal of Vegetable Crop Production*, **1**(2), 11–39.

[116] Skrabule, I. 2008 Comparison of potato clones developed and tested in organic and conventional growing conditions. *Agronomijas vēstis*, **11**, 147–154.

[117] Song, S., Lehne, P., Le, J., Ge, T. & Huang, D. 2009 Yield, fruit quality and nitrogen uptake of organically and conventionally grown muskmelon with different inputs of nitrogen, phosphorus, and potassium. *Journal of Plant Nutrition*, **33**(1), 130–141.

[118] Stonehouse, D. P., Weise, S., Sheardown, T., Gill, R. & Swanton, C. 1996 A case study approach to comparing weed management strategies under alternative farming systems in Ontario. *Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie*, **44**(1), 81–99.

[119] Suja, G., Sundaresan, S., John, K. S., Sreekumar, J. & Misra, R. S. 2012 Higher yield, profit and soil quality from organic farming of elephant foot yam. *Agronomy for sustainable development*, **32**(3), 755–764.

[120] Swezey, S., Rider, J., Werner, M., Buchanan, M., Allison, J. & Gliessman, S. 1994 In Santa Cruz County, Granny Smith conversions to organic show early success. *California Agriculture*, **48**(6), 36–44.

[121] Swezey, S. L., Goldman, P., Bryer, J. & Nieto, D. 2007 Six-year comparison between organic, IPM and conventional cotton production systems in the Northern San Joaquin Valley, California. *Renewable Agriculture and Food Systems*, **22**(1), 30–40.

[122] Szafirowska, A. & Elkner, K. 2008 Yielding and fruit quality of three sweet pepper cultivars from organic and conventional cultivation. *Vegetable Crops Research Bulletin*, **69**(1), 135–143.

[123] Tamm, I., Tamm, Ü. & Ingver, A. 2009 Spring cereals performance in organic and conventional cultivation. *Agronomy Research*, **7**, 522–527.

[124] Teasdale, J. R., Coffman, C. B. & Mangum, R. W. 2007 Potential long-term benefits of no-tillage and organic cropping systems for grain production and soil improvement. *Agronomy Journal*, **99**(5), 1297–1305.

[125] Temple, S. R., Friedman, D. B., Somasco, O., Ferris, H., Scow, K. & Klonsky, K. 1994 An interdisciplinary, experiment station-based participatory comparison of alternative crop management systems for California's Sacramento Valley. *American Journal of Alternative Agriculture*, **9**, 64–64.

[126] Thorup-Kristensen, K., Dresbøll, D. B. & Kristensen, H. L. 2012 Crop yield, root growth, and nutrient dynamics in a conventional and three organic cropping systems with different levels of external inputs and n re-cycling through fertility building crops. *European Journal of Agronomy*, **37**(1), 66–82.

[127] Torstensson, G., Aronsson, H. & Bergström, L. 2006 Nutrient use efficiencies and leaching of organic and conventional cropping systems in Sweden. *Agronomy Journal*, **98**(3), 603–615.

[128] Treadwell, D. D., Creamer, N. G., Hoyt, G. D. & Schultheis, J. R. 2008 Nutrient management with cover crops and compost affects development and yield in organically managed sweetpotato systems. *HortScience*, **43**(5), 1423–1433.

[129] Valkila, J. 2009 Fair trade organic coffee production in Nicaragua: Sustainable development or a poverty trap? *Ecological Economics*, **68**(12), 3018–3025.

[130] Wang, G., Ngouajio, M., McGiffen, M. E. & Hutchinson, C. M. 2008 Summer cover crop and in-season management system affect growth and yield of lettuce and cantaloupe. *HortScience*, **43**(5), 1398–1403.

[131] Warman, P. R. & Havard, K. 1997 Yield, vitamin and mineral contents of organically and conventionally grown carrots and cabbage. *Agriculture, ecosystems & environment*, **61**(2), 155–162.

[132] Warman, P. R. & Havard, K. 1998 Yield, vitamin and mineral contents of organically and conventionally grown potatoes and sweet corn. *Agriculture, ecosystems & environment*, **68**(3), 207–216.

[133] Welsh, C., Tenuta, M., Flaten, D., Thiessen-Martens, J. & Entz, M. 2009 High yielding organic crop management decreases plant-available but not recalcitrant soil phosphorus. *Agronomy Journal*, **101**(5), 1027–1035.

[134] Wheeler, S. A. & Crisp, P. 2011 Going organic in viticulture: a case-study comparison in Clare Valley, South Australia. *Australasian Journal of Environmental Management*, **18**(3), 182–198.

[135] Wisconsin Integrated Cropping Systems Trial (WICST) 2007 WICST annual crop yields.

Table S1: Parameter posteriors for models without explanatory variables. μ is the true mean response ratio across years and studies, σ_α is the standard deviation of the distribution from which the study random effects are drawn; σ_η is the standard deviation of the distribution from which the within year random effects are drawn; CV_{σ_η} is the coefficient of variation of the gamma from which the study-specific within-year variance are drawn; and σ_β is the standard deviation of the distribution of random between year effects. Values of Rhat < 1.1 indicate convergence. Lower Deviance Information Criterion (DIC) indicates better model fit to the data.

Parameter	Posterior mean	Posterior standard deviation	95% CI	Rhat
Study random effect, DIC=1684.8				
μ	0.795	0.027	0.742 – 0.848	1.001
σ_α	0.341	0.026	0.294 – 0.396	1.001
Study and within year random effects, DIC= -565.9				
μ	0.788	0.021	0.749 – 0.829	1.001
σ_α	0.188	0.024	0.144 – 0.239	1.001
σ_η	0.312	0.011	0.291 – 0.333	1.001
Study and study-specific within year random effects, DIC= -618.0				
μ	0.808	0.019	0.771 – 0.845	1.001
σ_α	0.189	0.023	0.145 – 0.237	1.001
CV_{σ_η}	1.155	0.135	0.907 – 1.436	1.001
Study, study-specific within year, and between year random effects, DIC= -621.2				
μ	0.808	0.019	0.770 – 0.846	1.001
σ_α	0.186	0.024	0.142 – 0.234	1.001
CV_{σ_η}	1.157	0.136	0.907 – 1.440	1.001
σ_β	0.041	0.027	0.002 – 0.098	1.001

Table S2: Yield estimates compared with previous studies. Categories of explanatory variables are arranged from the smallest to the largest difference between organic and conventional yields. Bold categories indicate conventional yields are significantly larger than organic yields. Unbolded categories indicate organic and conventional yields are not significantly different. ≪ between two categories indicates the two categories are significantly different from each other. ∗ includes both plant and animal products. ∗∗ de Ponti et al. [2] did not report significance. ∗ ∗ ∗ Confidence intervals were calculated from the standard deviation and the number of yield comparisons reported by de Ponti et al. [2].

Variable	**Ponisio et al.**	**Seufert et al. [1]**	**de Ponti et al. [2]**	**Badgley et al. [20]**
Overall	80.8% ± 3.7%	75% ± 4%	80% ± 1% ∗ ∗∗	132% ± 1%∗
Development	**developing** **developed**	**developed** ≪ **developing**	developing ∗∗ developed	**developing** ≪ **developed**
Crop type	**fruits and nuts** **oilseed crops** **cereals** **vegetables** **roots & tubers**	fruits and nuts **oilseed crops** ≪ **cereals** **vegetables**	vegetables ∗∗ cereals roots & tubers oilseed crops fruits	NA
Crop species	apple oat tomato **soybean** **maize** ≪ **wheat** **barley** **potato**	**soybean** **maize** **tomato** **barley** **wheat**	soybean ∗∗ maize oat tomato wheat barley apple	NA
Legume	**legume** **non-legume**	legume ≪ **non-legume**	NA	NA
Plant-type	**perennial** **annual**	perennial ≪ **annual**	NA	NA

Table S3: The impact of management practices on the yield gap compared with previous studies. Categories of explanatory variables are arranged from the smallest to the largest difference between organic and conventional yields. Bold categories indicate conventional yields are significantly larger than organic yields. Unbolded categories indicate organic and conventional yields are not significantly different.≪ between two categories indicates the two categories are significantly different from each other.

Variable	**Ponisio et al.**	**Seufert et al. [1]**
Nitrogen input	**similar N input ≪** **more N organic ≪** **more N conventional**	**more organic ≪** **more conventional** **similar N input**
Poly/monoculture	**organic polyculture ≪** **both monoculture** **both polyculture**	**organic polyculture** **both monocultures** **both polyculture**
Rotations	**more organic ≪** **similar** **no rotations**	**more organic** **similar ≪** **no rotations**
Organic fertilizer type	**organic fertilizer** **animal** **mix** **plant**	**organic fertilizer** **mix** **animal** **plant**
Conventional system type	**low input ≪** **high input**	low input ≪ **high input**
Organic system type	**certified** transitional **biodynamic** **organic standards**	**certified** biodynamic ≪ **organic standards** **transition**
Best management practices	**no** **yes**	**yes ≪** **no**
Irrigation	**rain-fed** **irrigated**	**rain-fed ≪** **irrigated**
Soil pH	acidic **neutral ≪** **strongly acidic** **strongly alkaline**	**weak acidic to alkaline ≪** **strongly acidic ≪** **strongly alkaline**
Time since conversion	**0-3 years** **> 7 years** **4-7 years**	**> 7 years** **4-7 years ≪** **0-3 years**

Table S4: The impact of study quality indicators on the yield gap compared with previous studies. Categories of explanatory variables are arranged from the smallest to the largest difference between organic and conventional yields. Bold categories indicated they are significantly less than one. $\ll$ between two categories indicates the two categories are significantly different from each other.

Variable	Ponisio et al.	Seufert et al. [1]
Duration of study	$>$ **10 seasons** **3-5 season** **6-10 seasons** **1-2 seasons**	$>$ **10 seasons** **6-10 seasons** **3-5 seasons** **1-2 seasons**
Literature type	**journal** **grey**	**grey** $\ll$ **journal**
Comparability of system	**comparable** **not comparable**	**comparable** **not comparable**
Study type	**on-farm trial** **experimental** **survey**	**on-farm trial** survey **experimental**

Table S5: A list of the studies included in the meta-analysis including the crops with yield comparisons, the country the study was conducted in, whether the study was included in Badgley et al. [20], Seufert et al. [1] or de Ponti et al. [2], and the number of organic to conventional yield comparisons extracted from each study.

Study	Crop	Country	Comparisons	In Badgley	In Seufert	In de Ponti
[21]	apple	Brazil	2	no	no	yes
[22]	pepper	India	5	no	yes	no
[23]	spring wheat, maize, soybean	United States	40	no	no	yes
[24]	wheat	Switzerland	2	no	no	yes
[25]	endive	Croatia	2	no	no	yes
[26]	rice, pigeon pea	India	2	no	no	no
[27]	apple	Switzerland	4	no	yes	yes
[28]	barley	Switzerland	4	no	yes	no
[29]	cabbage	Switzerland	4	no	yes	no
[30]	sugar beet	Switzerland	4	no	yes	no
[31]	tomato	Brazil	2	no	no	yes
[32]	wheat	Czech Republic	6	no	no	yes
[33]	cotton	India	1	no	yes	no

[34]	bean, cauliflower, fennel, lettuce, melon, tomato	Italy	41	no	no	no
[35]	maize, soybean, wheat	United States	32	no	yes	no
[36]	spinach	Turkey	34	no	yes	no
[37]	maize, tomato	United States	6	no	yes	yes
[38]	maize, soybean	United States	4	no	no	no
[39]	melon, watermelon	Turkey	6	no	no	yes
[40]	tomato, wheat	Italy	2	no	no	no
[41]	maize, soybean	United States	2	yes	yes	no
[42]	maize, soybean	United States	4	no	no	no
[43]	rice	France	9	no	no	no
[44]	hazelnut	Turkey	1	no	yes	no
[45]	tomato, maize	United States	45	no	yes	yes
[46]	maize, soybean	United States	2	yes	yes	no

[47]	barley, wheat	Denmark	78	no	yes	yes
[48]	tomato	United States	1	yes	yes	no
[49]	maize	United States	3	no	yes	no
[50]	safflower	United States	1	no	no	no
[51]	maize	Greece	9	no	no	yes
[52]	flax	Canada	1	no	yes	yes
[53]	cotton, chili, wheat, soybean, sorghum, maize	India	12	no	yes	yes
[54]	soybean, wheat, maize	United States	38	no	yes	yes
[55]	pea grain	Bulgaria	1	no	no	
[56]	strawberry	United States	3	no	yes	yes
[57]	maize	United States	3	no	yes	no
[58]	wheat	India	12	no	yes	no
[59]	soybean, wheat	India	8	no	no	no
[60]	kiwi	New Zealand	2	no	no	no
[61]	apricot	Turkey	3	no	no	yes
[62]	fig	Turkey	3	no	no	no
[63]	raisins	Turkey	3	no	no	yes
[64]	strawberry	Canada	8	no	yes	no
[65]	maize, soybean, oat	United States	6	no	no	no

[66]	chard, pumpkin, tomato, bean	Spain	4	no	yes	no
[67]	rye	Germany	2	no	yes	no
[68]	sweet potato	United States	2	no	no	yes
[69]	wheat, oat, barely	Estonia	3	no	no	yes
[70]	potato	Estonia	2	no	yes	no
[71]	banana	Ecuador	1	no	yes	no
[72]	tomato	Taiwan	3	no	yes	yes
[73]	buckwheat	Czech Republic	9	no	no	no
[74]	wheat	Canada	1	no	yes	yes
[75]	barley, wheat	Sweden	2	no	yes	yes
[76]	wheat	Australia	2	no	yes	yes
[77]	wheat, oat, soybean, maize	United States	4	no	no	no
[78]	soybean	China	1	no	yes	no
[79]	maize	Iran	12	no	no	no
[80]	oat, wheat, barley	Norway	86	no	no	no
[81]	maize, soybean	United States	4	no	yes	no
[82]	cauliflower	Italy	20	no	no	yes
[83]	maize	United States	26	no	yes	no

[84]	maize, soybean	United States	3	no	yes	no
[85]	maize, tomato, wheat	United States	28	no	yes	no
[86]	coffee	Costa Rica	3	no	yes	yes
[87]	strawberry	Turkey	5	no	no	yes
[88]	potato, wheat	Switzerland	18	yes	yes	yes
[89]	potato	Italy	1	no	no	yes
[90]	soybean, maize, oat	United States	10	no	no	no
[91]	grapes	Italy	3	no	no	yes
[92]	coffee	Nicaragua	3	no	yes	
[93]	maize, tomato	United States	5	no	yes	no
[94]	sunflower	Italy	3	no	yes	yes
[95]	wheat	Italy	2	no	yes	no
[96]	potato	Portugal	2	no	no	yes
[97]	leek	Croatia	4	no	no	yes
[98]	apple	United States	4	no	no	yes
[99]	apple	United States	4	no	no	no
[100]	maize	Italy	3	no	yes	no
[101]	tomato	United States	6	no	yes	no
[102]	lettuce	Turkey	5	no	yes	yes
[103]	tomato	Turkey	5	no	no	no

[104]	maize, oat, soybean, alfalfa	United States	24	no	yes	yes
[105]	maize, soybean	United States	8	no	yes	no
[106]	peach	United States	2	no	no	no
[107]	carrot, beetroot, rye, potato	Germany	8	yes	yes	no
[108]	rye	Germany	2	no	yes	no
[109]	wheat	United States	1	no	yes	no
[110]	apple	United States	2	yes	yes	yes
[111]	tomato	Tunisia	12	no	yes	yes
[112]	sweet corn, cucumber, pepper	United States	3	no	yes	no
[113]	wheat	Australia	1	no	yes	no
[114]	maize, soybean	United States	82	no	no	yes
[115]	cabbage, onion, sweet corn, bean, tomato	Canada	5	no	yes	no
[116]	potato	Latvia	2	no	no	no
[117]	muskmelon	China	4	no	no	no
[118]	wheat	Canada	2	no	yes	yes

[119]	elephant foot yam	India	5	no	no	no
[120]	apple	United States	1	no	yes	no
[121]	cotton	United States	2	no	yes	no
[122]	pepper	Poland	3	no	no	yes
[123]	wheat, oat, barley	Estonia	9	no	no	yes
[124]	maize, wheat	United States	13	no	yes	yes
[125]	tomato, maize, safflower, bean	United States	8	no	no	no
[126]	rye, oat, onion, carrot, cabbage, lettuce	Denmark	72	no	no	no
[127]	oat, barley	Sweden	3	no	yes	no
[128]	sweet potato	United States	3	no	no	yes
[129]	coffee	Nicaragua	1	no	yes	no
[130]	lettuce	United States	2	no	yes	yes
[131]	cabbage, carrot	Canada	6	yes	yes	yes
[132]	sweet corn, potato	Canada	6	yes	yes	yes
[133]	flax, wheat	Canada	13	no	yes	yes
[134]	grapes	Australia	1	no	no	no

[135]	soybean, wheat, maize	United States	3	no	yes	no

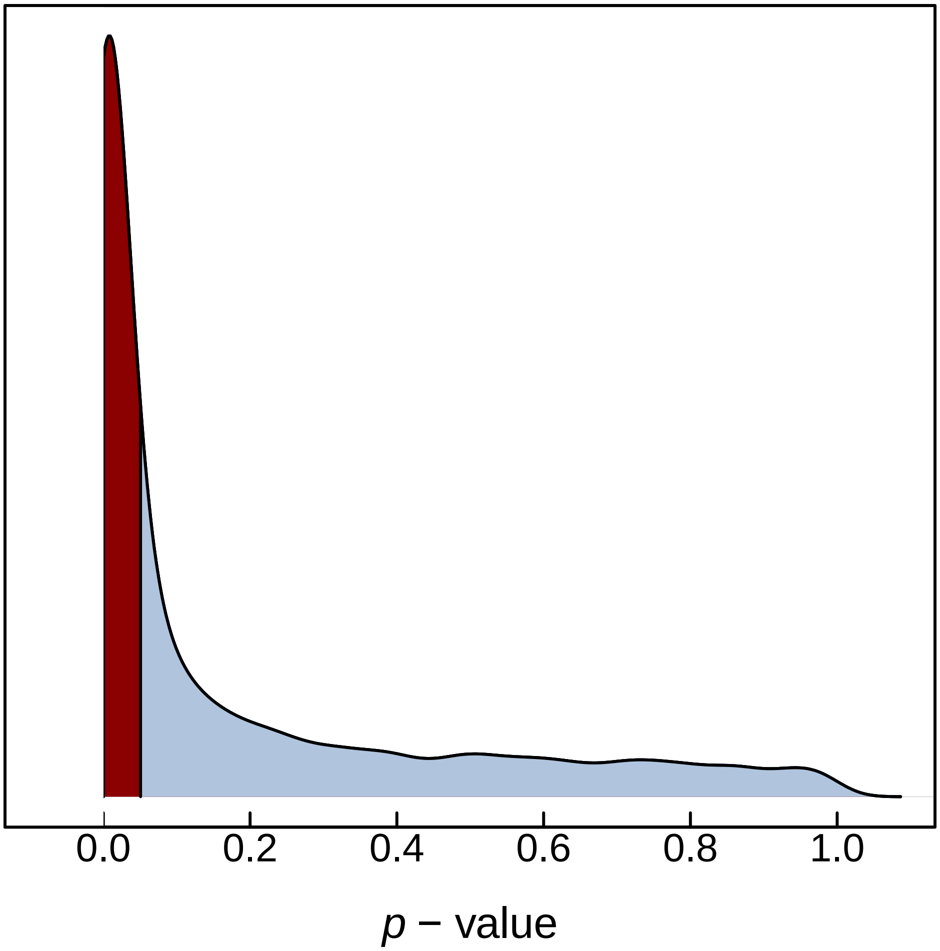

Figure S1: The distribution of *p*-values when the null hypothesis was forced to be true using the data and analysis type present in Seufert et al. [1]. If the analysis procedure was valid for these data, the distribution of P-values should be uniform between 0 and 1. Instead it is sharply shifted toward low P-values. In over 50% of simulations, the null hypothesis was rejected using a nominal Type I error rate of 0.05 (red region above).

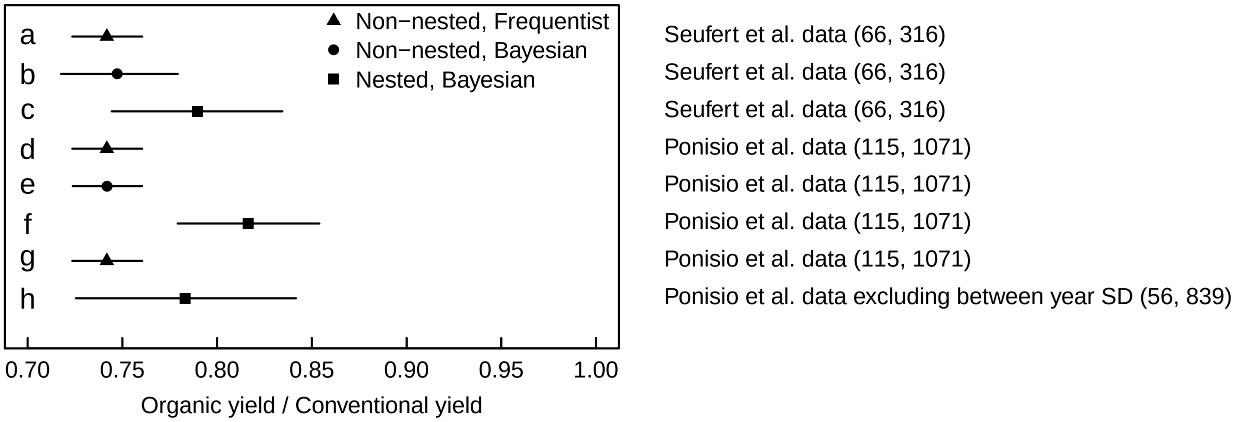

Figure S2: The effect of different models, data, and statistical paradigms on the organic to conventional yield ratio (from top to bottom): (a) Seufert et al.'s [1] non-nested analysis and data not accounting for pseudo-replication (organic to conventional ratio 75% ± 4%), (b) Seufert et al.'s [1] non-nested analysis and data conducted in a Bayesian framework to show the comparability in results between the two paradigms (75% ± 4%), (c) our nested analysis accounting for the hierarchy of the meta-dataset on the data published by Seufert et al. [1] in a Bayesian framework (79% ± 4.5%), (d) Seufert et al.'s [1] non-nested analysis on our larger dataset in a frequentist framework (74% ± 2%), (e) Seufert et al.'s [1] non-nested analysis on our larger dataset in a Bayesian framework (74% ± 2%), (f) our analysis and full meta-dataset (80.8% ± 3.7%), (g) an unweighted, non-nested analysis on our meta-dataset (74% ± 2%), and (h) our nested analysis on our meta-dataset excluding studies where the between year variation was used as an estimate of the observation-level sampling variance (78% ± 5.8%). The number of studies and organic to conventional comparisons are indicated in parentheses. We could not re-analyze de Ponti et al.'s [2] data because they did not extract the response ratio variances necessary for our model.

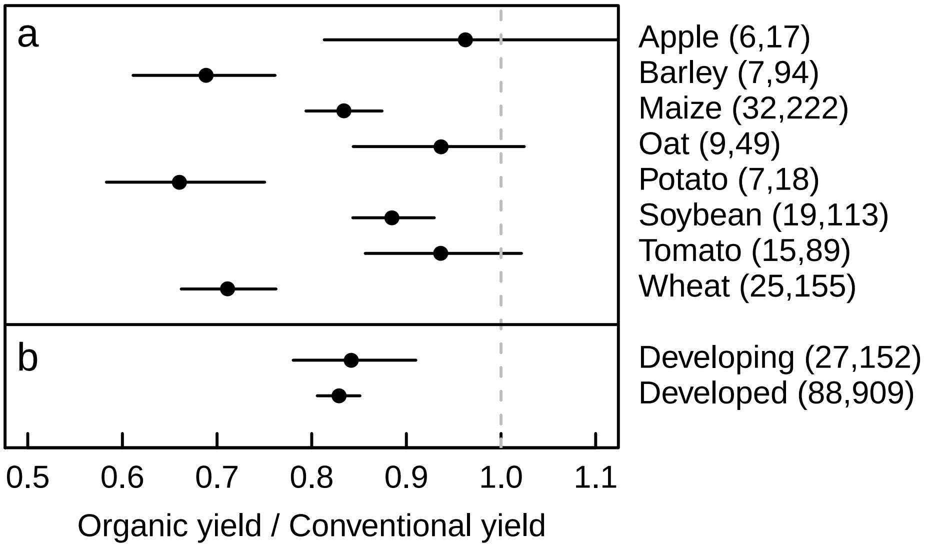

Figure S3: The influence of (a) crop species and (b) country development on the organic-to-conventional yield ratio. Values are mean effect sizes with 95% credible intervals. The number of studies and observations in each category are shown in parentheses. Only categories with at least 10 yield comparisons from greater than 5 studies are shown.

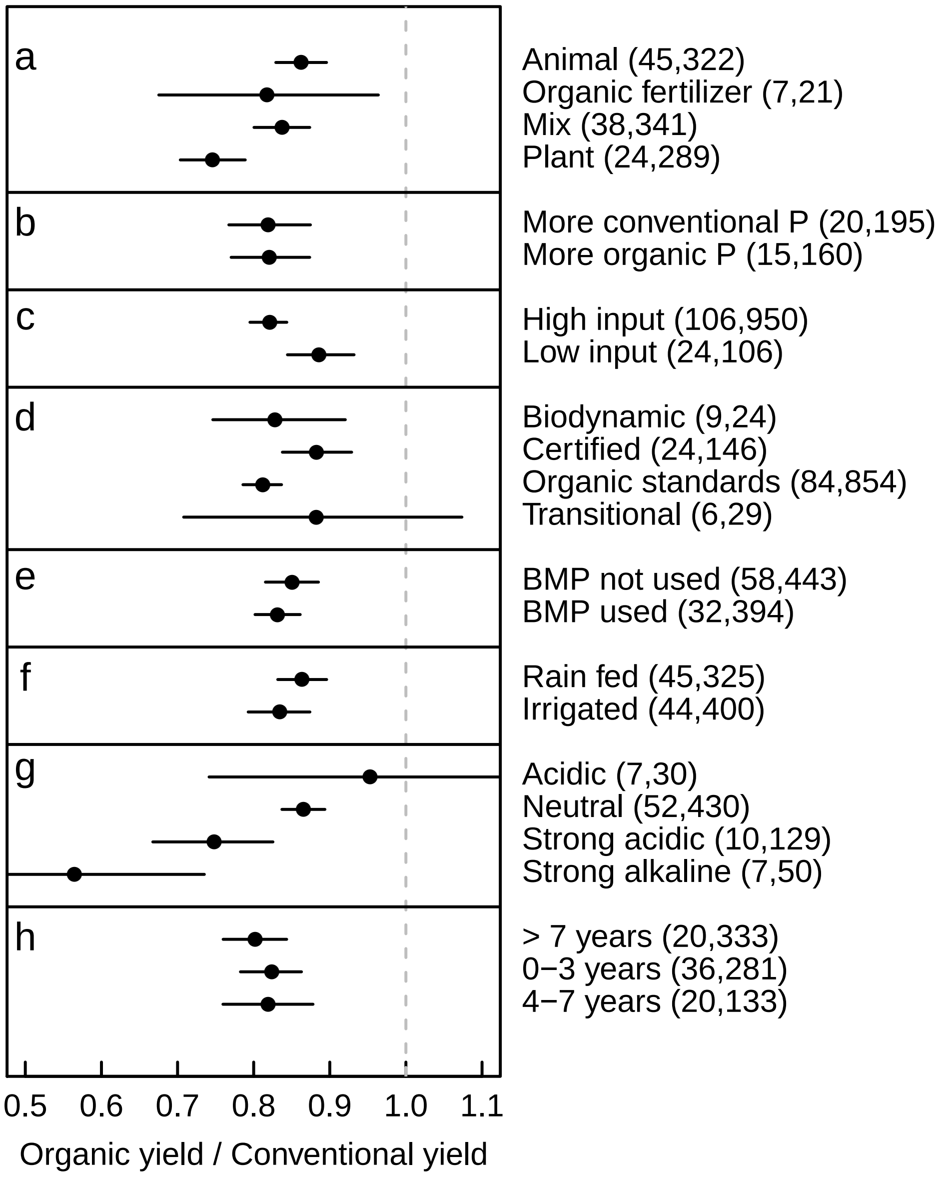

Figure S4: The influence of management practices on the organic-to-conventional yield ratio including (a) organic fertilizer type, (b) phosphorus input, (c) conventional system type, (d) organic system type, (e) the use of best management practices, (f) irrigation, (g) soil pH, and (h) years since conversion to organic agriculture. With respect to the low input conventional system type, the papers self-identified their system as "low input", but the crops generally received substantial external inputs. The inputs were generally only applied if needed (i.e., if a soil test suggested the soil was nitrogen deficient, fertilizer was applied), and not on a set schedule as is common in high-input conventional systems. Values are mean effect sizes with 95% credible intervals. The number of studies and observations in each category are shown in parentheses. Only categories with at least 10 yield comparisons from greater than 5 studies are shown.

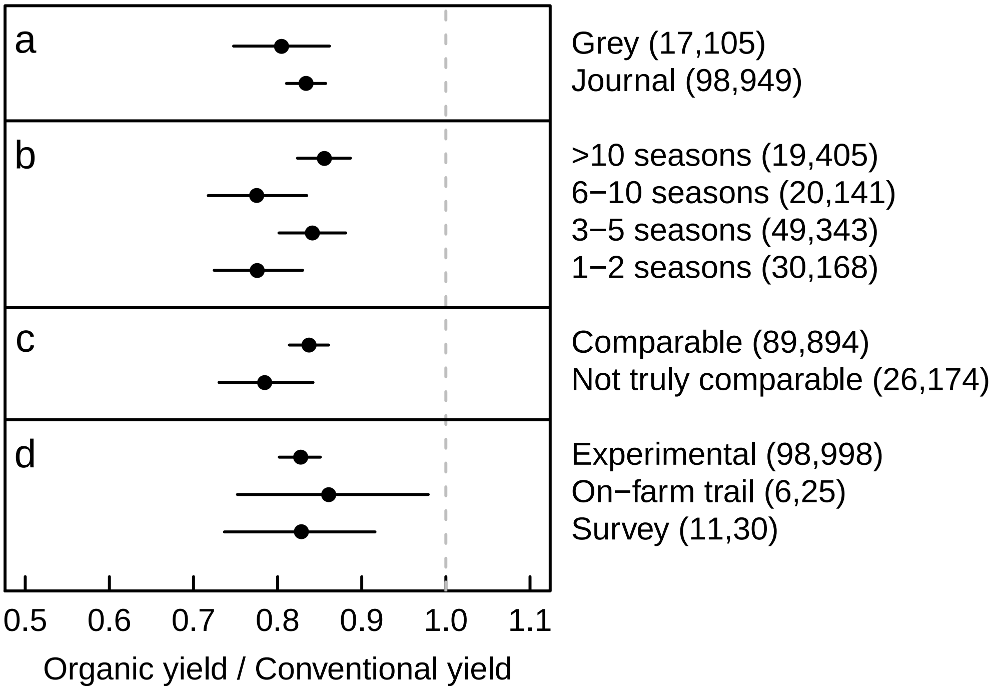

Figure S5: The sensitivity of the organic-to-conventional yield ratio to study quality factors including (a) publisher type, (b) duration of the study, (c) comparability of the organic and conventional treatments, and (d) the type of study. Values are mean effect sizes with 95% credible intervals. The number of studies and observations in each category are shown in parentheses. Only categories with at least 10 yield comparisons from greater than 5 studies are shown.

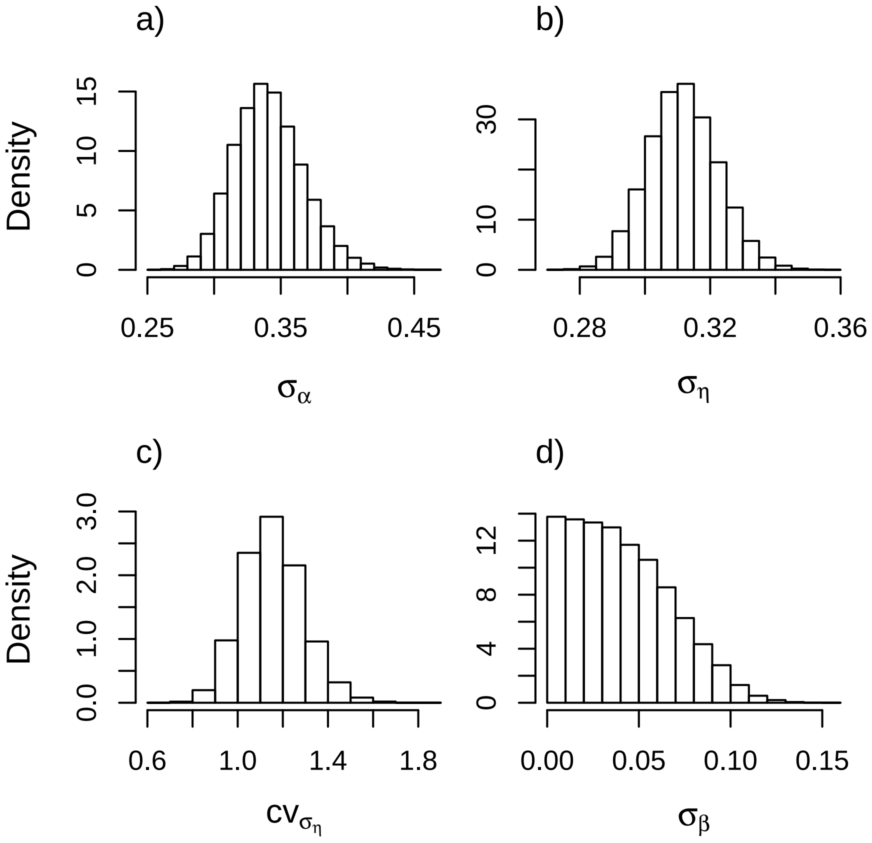

Figure S6: The posterior distributions for the random effect of a) study (σ_α); b) response ratios within a year (σ_η); c) response ratios within a year where the within year variance is study-specific, CV_{σ_η} is the coefficient of variation ($1/\sqrt{shape}$) of the gamma distribution (this model is most supported by the data); and d) between year (σ_β). Including a between-year variance term was not supported by the data (the posterior for (σ_β is not differentiated from zero).

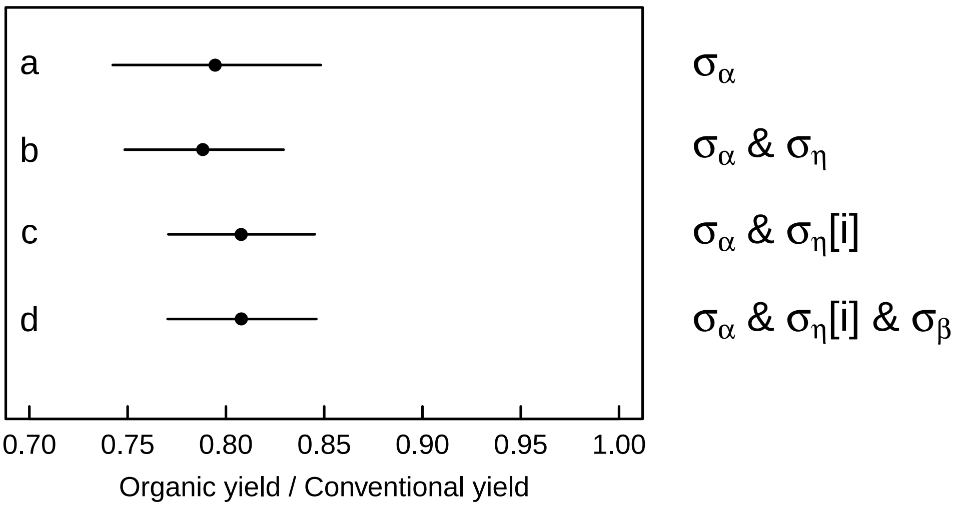

Figure S7: The sensitivity of the yield gap to including different levels of hierarchy in the model. The random effects included in the model are: a) study (σ_α); b) study and response ratios within a year (σ_η); c) study and response ratios within a year where the within year variance is study-specific ($\sigma_\eta[i]$) (this model is most supported by the data); and d) study, study-specific within-year variance, and between year (σ_β). Including a between-year variance term was not supported by the data (the posterior for (σ_β is not differentiated from zero). Values are the posterior mean with 95% credible intervals.

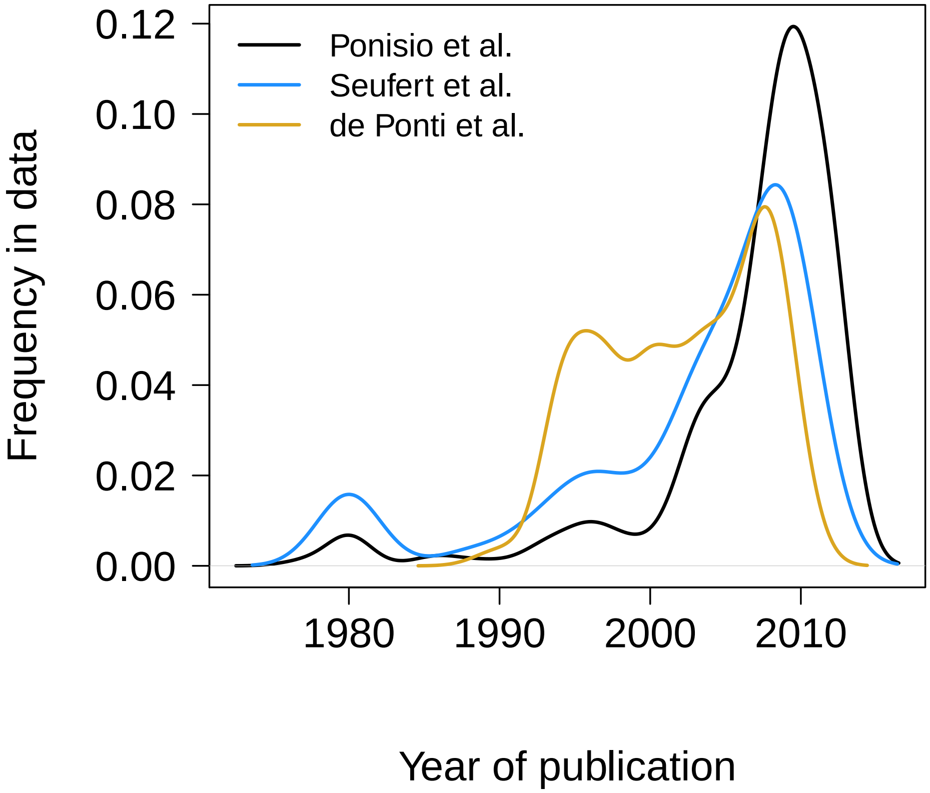

Figure S8: A comparison of the frequency of organic to conventional yield comparisons published in different years from our study, de Ponti et al. [2] and Seufert et al. [1].

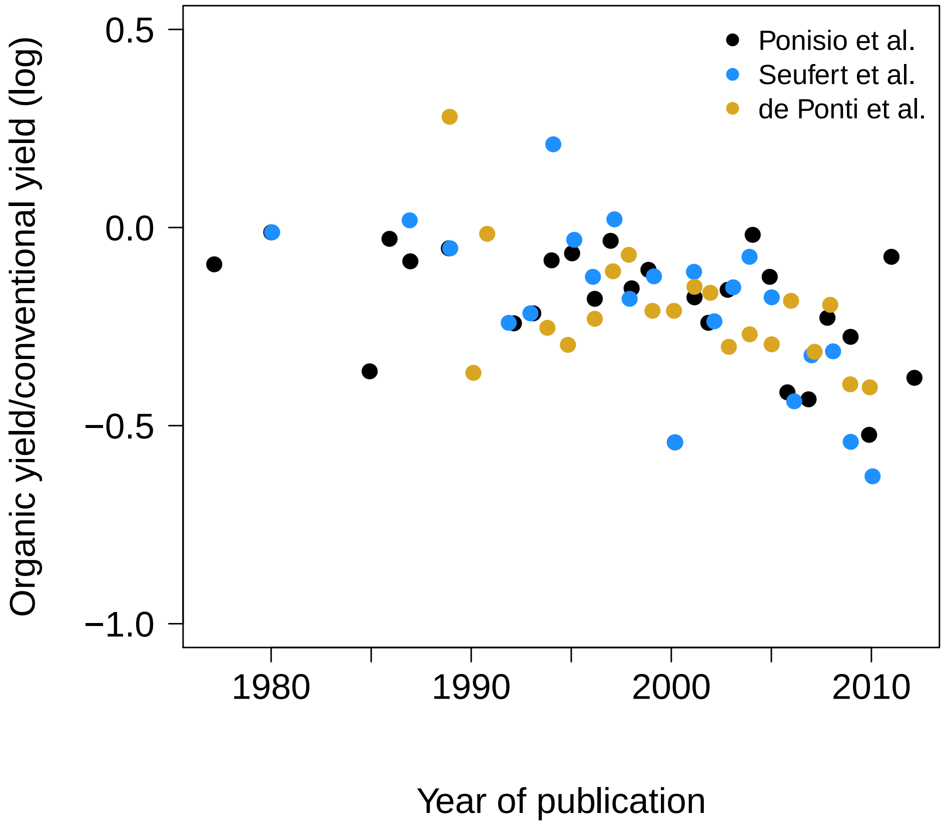

Figure S9: The mean organic to conventional yield ratio (log) for each year studies were published from our study, de Ponti et al. [2] and Seufert et al. [1]

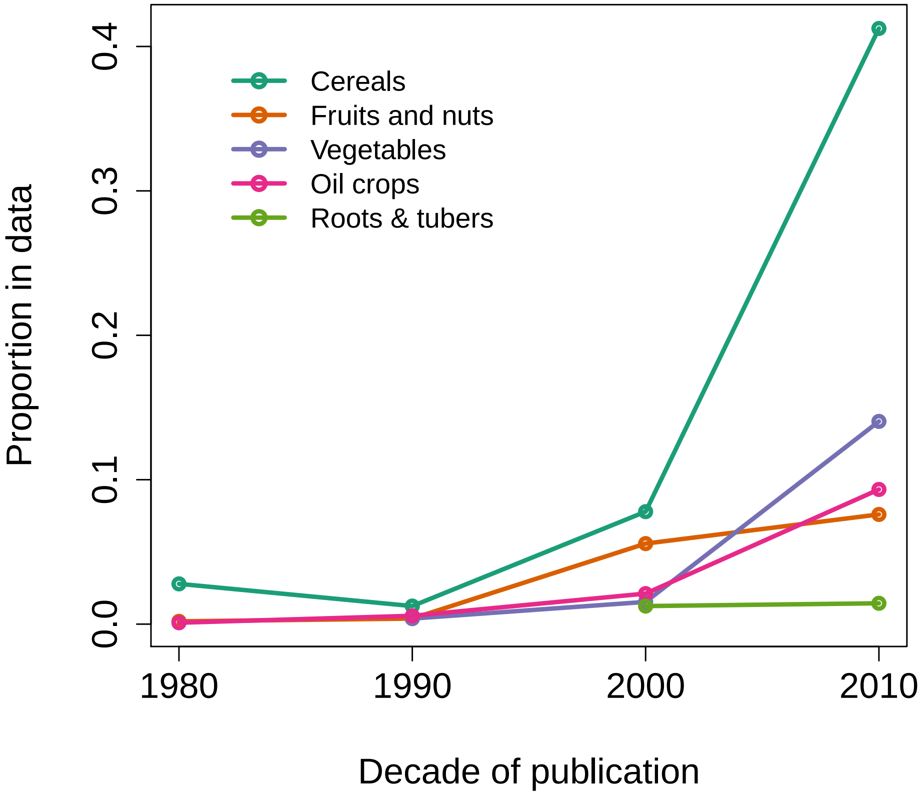

Figure S10: The proportion of observations of each crop species (binned by decade) in our meta-dataset.

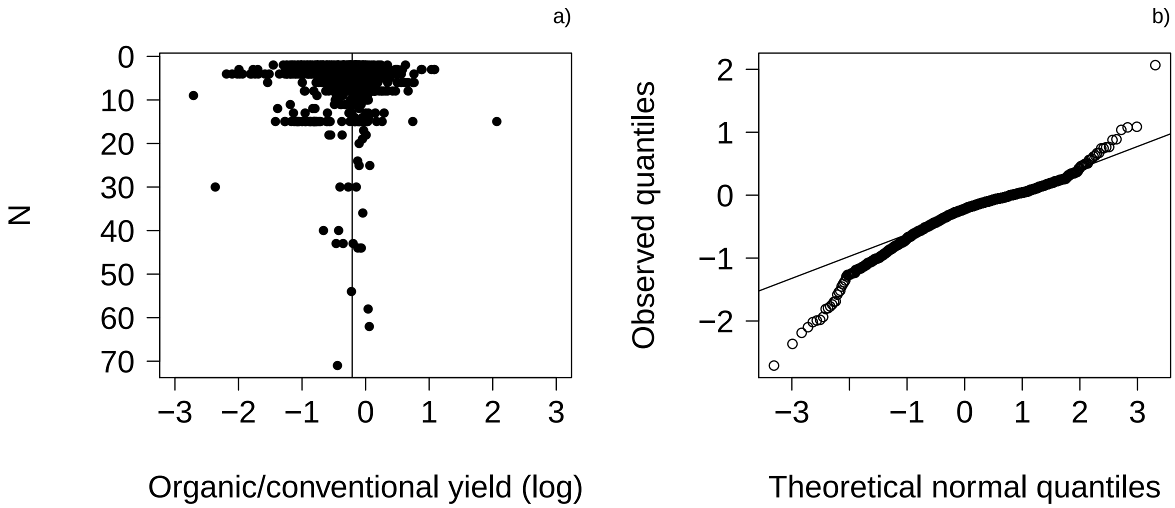

Figure S11: A funnel plot and QQ-plot of the response ratios in our study's meta-dataset.