

VOL. 7 NO.3 (77-120) JULI 2022

ISSN 2502-1710

AGRIMOR

**JURNAL AGRIBISNIS LAHAN
KERING**

DOI 1032938/ag.v7i3



About the Journal

AGRIMOR is an open access journal. All articles published through a peer review process. It is published four times a year, in January, April, July and October. Each time published at least five articles. International Standard Serial Number (ISSN) of AGRIMOR is 2502-1710.



Aims

Publishes primary research papers and review articles in areas of agribusiness.

Scope

AGRIMOR covers the field of agribusiness that studies management and socio-economic in the field of farming, agroindustry, marketing, information and communication systems with based on dry land agriculture including aspects of government policy.

Publisher

Faculty of Agriculture, Timor University.

Address

Eltari street, Km. 9, Village Sasi, Kefamenanu – Timor – NTT. Postcode 85613.
e-mail: savana.cendana@gmail.com
CP: +6281315023025

Editorial Team

Responsible Person	: Boanerges Putra Sipayung, S.P., M.P.
Editor in Chief	: Simon Juan Kune, S.P., M.P.
Managing Editor	: Boanerges Putra Sipayung, S.P., M.P.
Section Editor	: Boanerges Putra Sipayung, S.P., M.P.
Copyediting	: Umbu Joka, S.P., M.Si.
Layout Editor	: Ody Wolfrit Matoneng, S.P., M.Si.
Production Editor	: Philip Yosua Th. Dima, S.T.
Editorial Board	: Campina Illa Prihantini, S.E., M.Si., Sembilanbelas November Kolaka University, Indonesia. Dennis Mark Onuigbo, B.Agric., M.Sc., University of Nigeria, Nsukka. Ture Simamora, S.Pt., M.Si., Timor University, Indonesia. Dudi Septiadi, S.Pd., M.Si., University of Mataram, Indonesia. Umbu Joka, S.P., M.Si., Timor University, Indonesia. Achmad Subchiandi Maulana, SP., M.Si., Timor University, Indonesia. Laura Juita Pinem, S.Si., M.Si., Prima Indonesia University, Indonesia.

Website

www.savana-cendana.id

Indexed in

Tables of Contents

Household Consumption Patterns of Rice Farmer in the Village of Ketawang Gondang District Nganjuk Regency <i>Dwi Indah Febriyani, Sri Widayanti, Hamidah Hendrarini, Nisa Hafidhoh Fitriana</i>	77-85
Potential and Economic Feasibility of Peanut Cultivation in Rain Fedded Rice Field in Raddae Village, Wajo District <i>Syamsuri Syamsuri, Hasria Alang</i>	86-93
The Effect of Government Assistance on The Efficiency of Rice Farming in East Java <i>Rachman Hakim, Gazali Gazali</i>	94-101
Evaluating the Impact of Intervention Programs in Agricultural Sector <i>Muhammad Ibnu</i>	102-113
Social Capital Participatory Communication Model as Social Capital in Rural Development (Case Study of Noeltoko Village, West Miomaffo District, TTU Regency) <i>Ody Wolfrit Matoneng, Marsianus Falo</i>	114-120

Evaluasi Dampak Program Intervensi di Sektor Pertanian

Muhammad Ibnu

Fakultas Pertanian, Universitas Lampung, Bandar Lampung Indonesia

*Correspondence author : muhammad.ibnu@fp.unila.ac.id

Article Info

Article history:

Received 26 June 2022

Received in revised form 10 July 2022

Accepted 27 July 2022

DOI:

<https://doi.org/10.32938/aq.v7i3.1775>

Keywords:

Agricultural sector
Counterfactual
Impact evaluation
Intervention program
Source of bias

Abstrak

Interventions in the agricultural sector need to be evaluated whether they are really successful and have the expected impact on beneficiaries, such as farmers. However, the definition of impact evaluation has varied considerably in the last two decades. Some empirical studies are also still biased in evaluating the impact, and this is related to the method used. This research has three objectives. First, to review the definition of impact evaluation according to the literature. Second, to identify sources of bias in impact evaluation studies. Third, to review empirical studies on impact evaluation, which have different results, especially from the methodology used. Empirical studies that serve as 'cases' in the agricultural sector are studies on the impact of coffee standards and certification. Thus, this research method is a theoretical literature review supported by empirical research results. The literature review reveals that one of the most common definitions of impact evaluation is that evaluation relates to counterfactual, which is a comparison between what actually happened and what would have happened in the absence of the intervention. Inappropriately designed counterfactuals are a major source of bias in impact evaluation, in addition to validity threats. Based on a review of empirical studies on the impact of coffee standards and certification, it is concluded that some studies have more reliable results than others. Studies with more reliable results use reliable counterfactuals, so as to minimize impact evaluation bias through an 'apple-to-apple' comparison between the intervention group and the control group.

1. Pendahuluan

Program intervensi (misalnya, program penyuluhan pertanian, program bantuan pupuk, bibit dan lain-lain) di sektor pertanian memiliki tujuan tertentu dan target (penerima manfaat) tertentu pula. Namun, tidak akan jelas apakah intervensi tersebut benar-benar berhasil dan memberikan dampak yang diharapkan kepada sang penerima manfaat, kecuali jika dampaknya dievaluasi. Evaluasi dampak (*impact evaluation*) jelas diperlukan paling tidak untuk tiga keperluan. Pertama, untuk membantu pembuat kebijakan memutuskan apakah program menghasilkan dampak yang diinginkan. Kedua, untuk mempromosikan akuntabilitas dalam alokasi sumber daya di seluruh program publik (pemerintah). Ketiga, untuk mengisi kesenjangan dalam memahami apa yang berhasil dan apa yang tidak, dan bagaimana perubahan terukur dalam kesejahteraan dikaitkan dengan intervensi yang dilakukan (Williams dan Kedir 2020). Oleh karena itu, evaluasi dampak harus menyediakan desain dan/atau metode evaluasi yang tepat untuk menganalisis apakah partisipan benar-benar mendapat manfaat dari program—dan bukan dari parameter atau penyebab lain yang memengaruhi hasil (Crato dan Paruolo 2018).

Penelitian ini menggunakan standar dan sertifikasi untuk kopi sebagai 'kasus' atau 'contoh' program intervensi di sektor pertanian. Alasannya, standar dan sertifikasi kopi merupakan program intervensi yang paling banyak dievaluasi dampaknya (Traldi, 2021). Standar dan sertifikasi dapat disebut sebagai program intervensi (oleh swasta) karena bertindak sebagai mekanisme yang digerakkan oleh pembeli (*buyer*) untuk mempengaruhi sistem produksi petani (*producer*) secara ekonomi, sosial, dan lingkungan (Grabs, 2020). Teori dan/atau asumsi intervensinya adalah, jika petani melakukan proses produksi kopi sesuai standar dan sertifikasi, maka hal ini akan meyakinkan konsumen untuk memberikan nilai tinggi (secara ekonomi/harga) pada karakteristik sosial dan lingkungan yang melekat pada produk kopi tersebut. Selanjutnya, karena konsumen bersedia membayar harga tinggi (yang disebut juga harga premium) untuk kopi bersertifikat, petani kopi menerima keuntungan finansial yang lebih tinggi dari aktivitas perdagangan (Meemken *et al.*, 2021).

Namun, studi-studi empiris yang mengevaluasi dampak standar dan sertifikasi bervariasi dalam hasilnya, dari 'positif' hingga 'terbatas' hingga 'negatif'. Studi-studi yang positif mengenai manfaat sertifikasi menyatakan bahwa petani bersertifikat memperoleh manfaat ekonomi berupa harga kopi yang lebih tinggi dan kepemilikan lahan yang lebih terjamin (Ranjan Jena dan Grote, 2017). Petani bersertifikat juga memperoleh peningkatan kualitas dan produktivitas (Ruben dan Zuniga, 2011), peningkatan pendidikan dan aset infrastruktur serta investasi moneter. Sertifikasi membawa keuntungan ekonomi dalam jangka pendek, sedangkan pemberdayaan dan peningkatan kapasitas akan terlihat dalam jangka panjang (Raynolds, 2004)

Studi-studi lain menyatakan bahwa standar dan sertifikasi memiliki dampak ekonomi yang terbatas. Valkila (2009), misalnya, berpendapat bahwa keuntungan ekonomi dari sertifikasi *Fairtrade* organik bersifat terbatas karena produksi pertanian organik yang rendah. Ruben dan Fort (2012) juga menemukan bahwa meskipun memberikan manfaat kepada petani dalam hal penguatan organisasi (efek tidak langsung), *Fairtrade* hanya memberikan pendapatan bersih yang terbatas kepada petani. Begitu pula, meskipun menemukan dampak positif dari sertifikasi, Bacon (2005) dan Bacon *et al.*, (2008) juga mengingatkan bahwa baik petani bersertifikat dan tidak bersertifikat hampir sama rentannya akibat penurunan harga kopi, terutama di akhir 1990-an dan awal 2000-an (Bacon, 2005). Pendapatan yang rendah, relokasi yang tinggi, dan kerawanan pangan merupakan masalah yang terus berlanjut terutama di antara petani kopi berskala kecil (Bacon *et al.*, 2008).

Studi-studi lain relatif pesimistis, bahkan cenderung negatif, terutama mengenai dampak ekonomi sertifikasi. Kilian *et al.* (2004) berpendapat bahwa tidak semua skema sertifikasi sesuai untuk petani kecil. Skema *Fairtrade* memang baik untuk petani kecil apalagi jika menawarkan harga premium. Namun, harga premium sering kali tidak dapat mengimbangi beban biaya petani kecil atas tuntutan praktik lingkungan dan sosial yang tinggi (Glasbergen, 2018). Skema sertifikasi tunggal tidak akan membawa dampak yang sukses terhadap lingkungan. Kombinasi sertifikasi organik, *Fairtrade*, dan tanaman naungan (*shade certification*) akan menghasilkan pendekatan konservasi yang lebih baik, tetapi akan meningkatkan biaya petani (Philpott *et al.*, 2007). Penelitian lain juga menemukan bahwa sebagian besar petani menghadapi hambatan untuk memasuki pasar *Fairtrade* organik yang dipromosikan sebagai perdagangan yang adil. Penyebabnya adalah biaya sertifikasi organik cukup tinggi tetapi hasil produksi dengan metode organik rendah (Calo dan Wise, 2005; Riedl, 2009; Maclellan, 2016). Dalam studi mereka, Dietz *et al.*, (2020) dan Lyngbaek dan Muschler (2001) juga menyatakan bahwa biaya sertifikasi meningkatkan biaya usaha tani. Biaya sertifikasi menyebabkan harga premium organik tidak dapat menutupi total biaya yang dikeluarkan oleh petani bersertifikat (Lyngbaek dan Muschler 2001). Dengan demikian, pendapatan bersih petani bersertifikat organik sebenarnya lebih rendah dibandingkan petani konvensional.

Berdasarkan uraian di atas, penelitian ini memiliki tiga tujuan. Pertama, meninjau kembali definisi evaluasi dampak menurut literatur. Kedua, mengidentifikasi sumber bias dalam studi evaluasi dampak. Ketiga, meninjau kembali studi-studi tentang evaluasi dampak, yang memiliki hasil berbeda, terutama dari metodologi yang digunakan. Terdapat dua alasan utama melakukan penelitian ini. Pertama, banyak studi yang mempergunakan metode perhitungan evaluasi dampak yang bias, sehingga hasil evaluasinya akan bias pula. Kedua, penelitian ini berkontribusi pada literatur evaluasi dampak di sektor pertanian, yaitu dengan menjernihkan masalah perdebatan tentang dampak aktual standar dan sertifikasi, terutama dari sudut pandang definisi dampak dan metode yang digunakan.

2. Metode Penelitian

Penelitian ini menggunakan metode kualitatif, yaitu mengkaji (*review*) literatur yang bersifat teoritis dan empiris. Tujuan pertama dan kedua penelitian dicapai dengan telaah literatur yang bersifat teoritis, sedangkan tujuan ketiga dicapai melalui kajian studi-studi empiris. Kajian literatur yang bersifat teoritis berfungsi untuk menelaah definisi dampak yang beraneka ragam dan mengerucutkannya ke dalam definisi yang paling umum (Khandker *et al.*, 2009; Gertler *et al.*, 2016; Crato dan Paruolo, 2018). Kajian literatur yang bersifat teoretis juga berguna untuk mengidentifikasi sumber bias yang terjadi dalam evaluasi dampak. Kajian studi-studi empiris berguna sebagai 'kasus' atau 'contoh' dari penelitian evaluasi dampak. Studi-studi empiris yang dikaji mengevaluasi dampak standar dan sertifikasi di sektor pertanian, yaitu kopi, dengan metode kuantitatif untuk menguji efek kausal sertifikasi (Bacon, 2005; Bacon *et al.*, 2008; Ruben dan Fort, 2012; Van Rijsbergen *et al.*, 2016; Ranjan Jena dan Grote, 2017). Penelitian ini fokus meninjau metode kuantitatif yang digunakan oleh penelitian-penelitian empiris tersebut, apakah menggunakan metode evaluasi yang andal atau tidak. Hasil tinjauan mengenai metode ini sebagian besar bisa menjelaskan mengapa hasil-hasil studi-studi empiris terkait dampak standar dan sertifikasi masih berdebat satu sama lain.

3. Hasil Dan Pembahasan

Bagian ini menampilkan hasil kajian mengenai definisi evaluasi dampak menurut literatur, dilanjutkan dengan identifikasi sumber bias dalam evaluasi dampak, dan tinjauan tentang studi-studi evaluasi dampak di sektor pertanian, khususnya kopi.

Definisi evaluasi dampak menurut literatur

Evaluasi dampak telah mengalami perubahan definisi selama dua puluh tahun terakhir, sehingga metode evaluasinya juga berkembang mengikuti perubahan definisi tersebut (Khandker *et al.*, 2009; Gertler *et al.*, 2016; Crato dan Paruolo, 2018). Sebagai contoh, Inisiatif Internasional untuk Evaluasi Dampak (*the International Initiative for Impact Evaluation*), Bank Dunia (*the World Bank*), Badan Perlindungan Lingkungan AS (*the US Environmental Protection Agency*), Kompetisi Usaha Sosial Global (*the Global Social Venture Competition*), dan Organisasi Kerja Sama dan Pembangunan Ekonomi (*the Organization for Economic Co-operation and Development*), melakukan evaluasi dampak berdasarkan definisi mereka masing-masing.

Inisiatif Internasional untuk Evaluasi Dampak mendefinisikan evaluasi dampak sebagai 'Analisis yang mengukur perubahan bersih dalam hasil untuk sekelompok orang tertentu yang dapat dikaitkan dengan program tertentu menggunakan metodologi terbaik yang tersedia, layak dan sesuai dengan pertanyaan evaluasi yang sedang diselidiki dan untuk konteks tertentu' (*International Initiative for Impact Evaluation (3ie)*, 2008). Menurut Bank Dunia 'Evaluasi dampak membandingkan hasil suatu program dengan kontra faktual untuk menunjukkan apa yang akan terjadi pada penerima manfaat tanpa program. Tidak seperti bentuk evaluasi lainnya, evaluasi dampak menggunakan atribut perubahan yang diamati dalam hasil dari program yang dievaluasi dengan mengikuti desain eksperimen dan kuasi-eksperimen' (Williams, 2020). Menurut Badan Perlindungan Lingkungan AS evaluasi dampak adalah 'Jenis evaluasi yang menentukan perbedaan yang dibuat oleh suatu program. Ini dicapai dengan membandingkan hasil yang diamati dengan perkiraan apa yang akan terjadi jika tidak ada program tersebut' (*US Environmental Protection Agency*, 2022). Sedangkan menurut organisasi Kompetisi Usaha Sosial Global, evaluasi dampak dilihat sebagai penilaian 'Porsi dari hasil total yang terjadi sebagai akibat dari aktivitas perusahaan, di atas dan di luar apa yang akan terjadi atau solusi alternatif terbaik berikutnya. Dalam ilmu sosial, seseorang membutuhkan apa yang disebut sebagai kontra faktual, agar sebanding dengan kondisi eksperimen guna membedakan efek variabel dependen dari antara semua faktor lain yang dapat menyebabkan perubahan' (Schrefler, 2016). Menurut Organisasi Kerja Sama dan Pembangunan Ekonomi, evaluasi dampak adalah 'Penilaian tentang bagaimana intervensi yang dievaluasi memengaruhi hasil, apakah dampaknya disengaja atau tidak. Analisis dampak yang tepat membutuhkan kontra faktual tentang apa yang akan terjadi jika tidak ada intervensi' (*the Organization for Economic Co-operation and Development*, 2022)

Dengan begitu banyaknya definisi dampak dalam dua dekade terakhir, peneliti akan kesulitan memilih definisi evaluasi dampak yang sesuai. Namun, definisi evaluasi dampak dapat dirangkum dalam empat definisi yang paling umum. Pertama, evaluasi dampak merupakan evaluasi yang melihat dampak intervensi pada hasil akhir (*outcomes*) kesejahteraan, bukan hanya pada keluaran (*outputs*) proyek, atau evaluasi proses yang berfokus pada implementasi (McGill *et al.*, 2021). Kedua, evaluasi dampak adalah evaluasi yang terkait dengan penetapan kontra faktual, yaitu perbedaan yang dibuat oleh suatu program intervensi dibandingkan dengan bila program intervensi tersebut tidak ada (Khandker *et al.*, 2009; Gertler *et al.*, 2016; Crato dan Paruolo, 2018). Ketiga, evaluasi dampak merupakan evaluasi yang dilakukan beberapa waktu (lima sampai sepuluh tahun) setelah program intervensi selesai sehingga memberikan waktu bagi munculnya dampak (Dietz, 2021). Keempat, evaluasi dampak adalah evaluasi yang mempertimbangkan semua intervensi dalam sektor atau wilayah geografis tertentu (Mancini dan Sala, 2018).

Sumber bias dalam evaluasi dampak: masalah kontra faktual

Penelitian ini mengikuti salah satu definisi yang paling umum, yaitu evaluasi dampak didefinisikan sebagai 'kesenjangan atau perbedaan antara hasil aktual dan hasil kontra faktual' (Khandker *et al.*, 2009; Gertler *et al.*, 2016; Crato dan Paruolo, 2018). Berdasarkan definisi ini, evaluasi dampak didesain untuk menjawab pertanyaan: bagaimana hasil (misalnya produksi kopi petani) telah berubah jika standar dan sertifikasi tidak ada? Pertanyaan ini terkait dengan analisis kontra faktual (*counterfactual*), yaitu perbandingan antara apa yang sebenarnya terjadi (fakta) dan apa yang akan terjadi jika tidak ada intervensi (kontra faktual) (Khandker *et al.*, 2009; Gertler *et al.*, 2016; Crato dan Paruolo, 2018). Prinsip desain kontra faktual bertujuan meminimalkan bias dengan menghilangkan atau meminimalkan adanya kemungkinan parameter (penyebab dampak) yang lain selain dari parameter yang diamati (Khandker *et al.*, 2009; Gertler *et al.*, 2016; Crato dan Paruolo, 2018).

Masalah pertama terkait kontra faktual yaitu kontra faktual tidak dapat diamati secara langsung sehingga menyulitkan dalam evaluasi dampak (Khandker *et al.*, 2009). Dengan demikian, tantangan sebenarnya dari evaluasi dampak adalah membentuk kelompok perbandingan yang 'meyakinkan' dan rasional untuk responden (Khandker *et al.*, 2009; Gertler *et al.*, 2016; Crato dan Paruolo, 2018). Idealnya, peneliti ingin membandingkan kelompok responden

yang sama ketika dengan dan tanpa intervensi. Dengan kata lain, jika suatu kelompok responden diberikan perlakuan (intervensi), dampak perlakuan tersebut dapat diukur dengan mengevaluasi kelompok responden yang sama ketika kelompok tersebut tidak diberi perlakuan (sebagai kontrol). Namun, suatu kelompok responden tidak dapat berada dalam posisi sedang diberi perlakuan dan posisi sebagai kontrol pada saat yang sama. Dengan kata lain, responden tidak dapat memiliki 'dua keberadaan dalam waktu yang bersamaan' (Khandker *et al.*, 2009).

Masalah kedua terkait kontra faktual yaitu penggunaan metode yang tidak bisa diandalkan (*unreliable*) dalam mendesain kontra faktual. Literatur memberikan contoh dua metode umum yang berpotensi bias (tidak bisa diandalkan) dalam mendesain kontra faktual, yaitu metode 'dengan-dan-tanpa perbandingan' (*with-and-without comparison*) dan metode 'perbandingan sebelum-dan-sesudah' (*before-and-after comparison*) (Khandker *et al.*, 2009; Gertler *et al.*, 2016; Crato dan Paruolo, 2018).

Metode 'dengan-dan-tanpa perbandingan' (*with-and-without comparison*) menggunakan kelompok non-partisipan sebagai kelompok kontrol, sehingga hasil non-partisipan merupakan hasil kontra faktual. Metode ini memiliki asumsi implisit bahwa jika petani bersertifikat belum bergabung dalam standar dan sertifikasi, hasil rata-rata mereka akan sama dengan petani yang tidak bersertifikat (Blackman dan Rivera, 2011). Namun, metode ini tidak bisa diandalkan karena hanya mendesain kontra faktual palsu dan rentan menyebabkan bias seleksi, yaitu bias yang terjadi ketika membandingkan dua identitas atau kelompok. Bias seleksi sering terjadi ketika suatu entitas atau kelompok memilih sendiri atau dipilih oleh pihak sertifikasi untuk dimasukkan ke dalam program sertifikasi (Khandker *et al.*, 2009; Gertler *et al.*, 2016; Crato dan Paruolo, 2018). Blackman dan Rivera (2011) memberikan contoh sebuah penelitian yang mengevaluasi dampak sertifikasi kopi organik terhadap erosi tanah. Dalam studi ini, erosi tanah di lahan petani non-sertifikasi berfungsi sebagai hasil kontra faktual. Dampak sertifikasi kopi organik diukur sebagai "perbedaan rata-rata ukuran erosi tanah antara rumah tangga petani sertifikasi dan petani non-sertifikasi" (Blackman dan Rivera, 2011). Namun, ada kemungkinan beberapa petani kopi telah mengadopsi konservasi tanah sebelum mereka terlibat dalam sertifikasi. Mereka secara sukarela bergabung dengan sertifikasi karena mereka tidak perlu berinvestasi dalam persyaratan konservasi tambahan untuk memenuhi standar sertifikasi. Oleh karena itu, jika hasil evaluasi dampak menunjukkan bahwa sertifikasi mengurangi erosi tanah, temuan ini adalah bias karena petani sudah melakukan konservasi tanah sebelum berpartisipasi dalam sertifikasi.

Metode 'perbandingan sebelum-dan-sesudah' (*before-and-after comparison*) disebut juga "metode refleksif" dampak, dimana hasil sebelum berpartisipasi dalam program intervensi berfungsi sebagai hasil kontrol. Metode ini berguna dalam evaluasi kebijakan nasional ketika seluruh populasi berpartisipasi dan tidak diperlukan kelompok kontrol. Namun, metode ini sangat mungkin bermasalah karena beberapa efek eksternal lainnya mungkin mempengaruhi dampak program intervensi selama proses implementasi. Sebagai contoh, ada sebuah studi tentang dampak sosial ekonomi dari sertifikasi *Fairtrade* terhadap petani kopi. Studi tersebut mengevaluasi dampak ekonomi sertifikasi berdasarkan perbedaan pendapatan rata-rata rumah tangga sebelum sertifikasi dan sesudah sertifikasi (Blackman dan Rivera, 2011). Studi tersebut menggunakan pendapatan rumah tangga petani bersertifikat sebelum sertifikasi sebagai kontra faktual. Katakanlah bahwa pendapatan rumah tangga sesudah sertifikasi lebih tinggi daripada sebelum sertifikasi, sehingga peneliti menyimpulkan bahwa sertifikasi meningkatkan pendapatan rata-rata rumah tangga petani. Namun, kesimpulan peneliti tersebut mungkin bias, karena mungkin terdapat faktor-faktor yang tidak terkait dengan sertifikasi yang mempengaruhi peningkatan pendapatan rumah tangga sesudah sertifikasi. Faktor-faktor tersebut dapat berupa peningkatan harga kopi internasional, perbaikan penanganan pasca panen, perbaikan sistem perdagangan, peningkatan peran koperasi, kondisi iklim yang mendukung, dan lain-lain (Blackman dan Rivera, 2011).

Sumber bias dalam evaluasi dampak: ancaman validitas

Semua desain evaluasi dampak berpotensi bias karena masalah validitas internal, validitas eksternal, dan validitas konstruk ((Khandker *et al.*, 2009; Gertler *et al.*, 2016; Crato dan Paruolo, 2018). Validitas internal mengacu pada 'apakah mungkin untuk menarik kesimpulan secara valid dan bahwa perbedaan pada variabel dependen adalah benar karena variabel penjelas yang diamati'. Validitas eksternal mengacu pada 'apakah hasil eksperimen atau eksperimen semu dapat digeneralisasikan ke konteks lain.' Validitas konstruk mengacu pada 'apakah peneliti benar-benar memahami sifat program intervensi yang dievaluasi'(Khandker *et al.*, 2009; Gertler *et al.*, 2016; Crato dan Paruolo, 2018).

Ancaman validitas internal terkait dengan adanya pembaur (*confounding*) dan bias seleksi. *Confounding* mengacu pada hubungan palsu di mana variabel eksternal, selain variabel yang

diamati, mempengaruhi hasil program intervensi. Bias seleksi pada umumnya terjadi ketika partisipan dipilih secara non-acak dari populasi yang memenuhi syarat untuk bergabung dalam program intervensi. Bias seleksi dibagi menjadi dua, yaitu bias seleksi mandiri (*self-selection bias*) dan bias seleksi program endogen (*endogenous program selection bias*). Bias seleksi mandiri terjadi ketika partisipan yang telah memiliki karakteristik yang sesuai dengan suatu program, secara sukarela berpartisipasi atau sengaja dipilih untuk mengikuti program tersebut. Bias seleksi program endogen terjadi ketika partisipan dipilih untuk berpartisipasi karena mereka diharapkan memperoleh manfaat lebih dari program tersebut (Khandker et al., 2009; Gertler et al., 2016; Crato dan Paruolo, 2018).

Ancaman validitas eksternal bersumber dari dua hal, yaitu populasi dan lingkungan. Ancaman validitas yang terkait dengan populasi disebabkan oleh kurangnya sampel yang diambil yang mewakili populasi tersebut. Ancaman yang terkait dengan lingkungan misalnya adalah efek Hawthorne (yaitu kesadaran partisipan untuk berpartisipasi dalam eksperimen) dan efek Rosenthal (yaitu kesadaran partisipan tentang harapan mereka) dan situasi tertentu (waktu, lokasi, ruang lingkup, tingkat pengukuran dan sebagainya) (Khandker et al., 2009; Gertler et al., 2016; Crato dan Paruolo, 2018).

Ancaman validitas konstruk terkait dengan kontaminasi (atau penularan) yang mempengaruhi 'dampak bersih' suatu program intervensi. Kontaminasi dapat berasal dari dua penyebab. Penyebab pertama kontaminasi adalah efek limpahan (*spillover effects*). Efek ini sering terjadi ketika kelompok intervensi dan kelompok kontrol secara geografis sangat dekat atau bertetangga, sehingga intervensi yang diberikan kepada partisipan sering juga sampai kepada non partisipan ((Khandker et al., 2009; Gertler et al., 2016; Crato dan Paruolo, 2018). Misalnya, informasi tentang penggunaan pupuk yang baik yang disampaikan kepada petani sertifikasi, dari mulut ke mulut, diterima juga oleh petani non-sertifikasi. Penyebab kontaminasi kedua adalah pengaruh tumpang tindih dari program intervensi lain selain yang sedang dievaluasi. Misalnya, sekelompok petani kopi mungkin ikut serta dalam sertifikasi *Fairtrade* dan sertifikasi *Utz* pada saat yang bersamaan. Oleh karena itu, dampak dari intervensi tunggal sampai batas tertentu akan lebih sulit untuk dievaluasi, dan peneliti harus sangat berhati-hati tentang hal ini (Blackman dan Rivera, 2011).

Tinjauan studi-studi tentang evaluasi dampak di sektor pertanian, khususnya kopi

Seperti dijelaskan sebelumnya pada bagian pendahuluan, dampak standar dan sertifikasi belum dapat digeneralisasi karena berbagai faktor yang berbeda, seperti lokasi, waktu penelitian, dan juga metode evaluasi dampak yang digunakan. Penelitian ini fokus melihat metode penelitian yang digunakan, terlepas dari hasil studi-studi tersebut tentang dampak standar dan sertifikasi, apakah positif atau terbatas atau negatif. Berdasarkan kajian tentang metode yang digunakan, sebagian studi empiris hasilnya lebih dapat dipercaya dibandingkan studi-studi yang lain. Studi-studi yang hasilnya lebih dapat dipercaya mempergunakan desain kontra faktual yang andal (*reliable*), sehingga mampu meminimalkan bias evaluasi dampak melalui perbandingan yang sesuai/proportional antara kelompok intervensi dan kelompok kontrol) (lihat Tabel 1).

Ada dua metode utama untuk mendesain kontra faktual yang andal (*reliable*), yaitu desain eksperimen (*experimental design*) dan desain eksperimen semu (*quasi-experimental design*) ((Khandker et al., 2009; Gertler et al., 2016; Crato dan Paruolo, 2018). Pada metode pertama, yaitu eksperimen, variasi dikontrol dengan cara 'mengendalikan bagaimana data dikumpulkan'. Cara pengumpulan data yang paling umum dalam eksperimen adalah melalui pengacakan (*randomization*), terutama untuk program yang memiliki satu fase dan variasi intervensi yang terbatas. Aspek yang terpenting dalam pengacakan (*randomization*) adalah setiap anggota populasi memiliki peluang yang sama untuk terpilih sebagai partisipan. Selain itu, partisipan (yaitu kelompok yang menerima intervensi) dan non-partisipan (yaitu, kelompok kontrol) harus identik secara statistik. Dengan cara ini, setiap perbedaan hasil antara kedua kelompok dapat dikaitkan dengan intervensi. Dalam eksperimen yang ideal, misalnya di laboratorium, perbedaan hasil (antara kelompok intervensi dan kelompok kontrol) dapat menghasilkan evaluasi mengenai dampak program yang sebenarnya (*actual impact*). Pengacakan (*randomization*) banyak digunakan dalam evaluasi dampak karena potensi biayanya yang lebih rendah daripada studi dampak berbasis survei lainnya. Selain itu, dengan pengacakan (*randomization*) peneliti tidak harus mengambil seluruh populasi sebagai partisipan; sampel yang jumlahnya memenuhi syarat, dalam banyak kasus, telah cukup untuk mewakili populasi (Khandker et al., 2009; Gertler et al., 2016; Crato dan Paruolo, 2018)(White, 2006; Ferraro, 2009; Khandker et al., 2009) .

Meskipun eksperimen secara acak berpotensi mengurangi bias seleksi, metode ini juga memiliki kelemahan untuk diterapkan pada evaluasi dampak program-program pembangunan. Paling tidak ada empat kelemahan desain eksperimen ((Khandker et al., 2009; Gertler et al., 2016; Crato dan Paruolo, 2018). Kelemahan pertama adalah kurangnya kesempatan untuk melakukan

pengacakan. Dalam kasus program intervensi dalam pembangunan, sebagian besar kejadian adalah masa lalu (*ex-post*) sehingga peluang peneliti untuk mengontrol data yang dikumpulkan dengan pengacakan telah berlalu. Kelemahan kedua adalah pengacakan mungkin sulit diterapkan pada intervensi yang kompleks. Kenyataannya, banyak program pembangunan yang begitu kompleks sehingga membutuhkan banyak perbandingan untuk mengatasi heterogenitasnya. Kelemahan ketiga adalah peneliti perlu mempertahankan kontrol dalam percobaan. Dalam beberapa kasus, hal ini justru mengarah pada masalah bias seleksi. Beberapa calon partisipan mungkin tidak ingin berpartisipasi sehingga peneliti melakukan intervensi pada partisipan seperti melobi mereka atau memberikan semacam insentif (yaitu, hadiah atau uang). Akibatnya, kemurnian desain pengacakan terkontaminasi. Dengan demikian pengacakan, sebagai elemen penting dari desain eksperimen, bukanlah obat mujarab untuk semua masalah terkait bias seleksi. Kelemahan ke empat adalah banyak desain eksperimen tidak layak karena alasan politik, keuangan, hukum, praktis, atau etis. Misalnya, di bidang kesehatan, tidak mungkin dan tidak etis untuk melakukan desain eksperimen untuk mengobati anak-anak yang terkena malaria, sementara pada saat yang sama tidak melakukan apa pun kepada anak-anak lain dengan alasan pembentukan kelompok kontrol (Schweizer *et al.*, 2016).

Metode kedua, sebagai alternatif desain eksperimen, adalah dengan menggunakan desain eksperimen semu (*quasi-experimental design*) (Khandker *et al.*, 2009; Gertler *et al.*, 2016; Crato dan Paruolo, 2018). Metode eksperimen semu ini dapat digunakan untuk mengatasi keterbatasan desain eksperimen dalam mengungkapkan korelasi kausal antara faktor-faktor dalam evaluasi dampak. Dengan metode ini, non-partisipan yang 'benar-benar sebanding' dibandingkan dengan partisipan. Perbandingan yang tepat (*apple-to-apple comparison*) menyeimbangkan semua faktor yang dapat diukur, seperti dalam eksperimen yang ideal. Pendekatan eksperimen semu fokus mengurangi bias seleksi dan mengoreksi perbandingan yang tidak tepat antara partisipan dan non-partisipan (Khandker *et al.*, 2009; Gertler *et al.*, 2016; Crato dan Paruolo, 2018). Setidaknya ada tiga cara untuk melakukan perbandingan yang tepat/adil. Pertama, non-partisipan yang sebanding dipilih secara acak dari komunitas. Kedua, menggunakan *database*, non-partisipan yang memiliki karakteristik yang sama dengan partisipan dipilih. Ketiga, non-partisipan dipilih dari sumber data lain tetapi memiliki karakteristik yang sama dengan partisipan dan memenuhi syarat untuk berpartisipasi dalam program (Schweizer *et al.*, 2016).

Dalam kasus sertifikasi kopi, petani sertifikasi dibandingkan dan/atau dicocokkan dengan petani non-sertifikasi. Sampel dari kedua kelompok harus diambil secara acak dan seimbang agar dapat dampak aktual sertifikasi pada petani dapat diukur (misalnya, dampak pada produksi, pendapatan dan pengeluaran, modal dan tabungan, serta pada sikap dan persepsi (Ruben dan Zuniga, 2011; Ruben dan Fort 2012; Van Rijsbergen *et al.*, 2016). Petani non-sertifikasi merupakan kelompok kontrol dan hasil (*output*) dari kelompok ini merupakan hasil (*output*) kontra faktual. Dengan perbandingan/pencocokan (*matching*) antara kelompok intervensi dan kelompok kontrol, metode kuasi-eksperimen atau eksperimen semu dapat membangun kontra faktual yang andal (*reliable*) dan biasanya menggunakan metodologi kuantitatif, seperti analisis regresi multivariat (Khandker *et al.*, 2009; Gertler *et al.*, 2016; Crato dan Paruolo, 2018). Namun, dalam desain kuasi eksperimen, peneliti harus berhati-hati agar tidak ada karakteristik tersembunyi atau tidak dapat diamati yang dapat mempengaruhi hasil evaluasi suatu program intervensi. Berbagai karakteristik, misalnya, pengetahuan dan keterampilan manajemen untuk kedua kelompok juga harus diupayakan agar serupa atau sebanding.

Tabel 1. Studi-studi dengan kontra faktual yang andal (*reliable*) dan studi-studi dengan kontra faktual yang tidak andal (*unreliable*)

Studi dengan kontra faktual yang andal	Studi dengan kontra faktual yang tidak andal
Lyngbaek dan Muschler (2001) melakukan studi tentang "Produktivitas dan profitabilitas pertanian kopi organik multistrata versus konvensional di Kosta Rika". Mereka membandingkan antara pertanian organik dan pertanian konvensional, terutama dalam hal output dan profitabilitas. Penulis membandingkan sepuluh pasang (<i>paired comparison</i>) pertanian organik dan konvensional. Pertanian konvensional bertindak sebagai kelompok kontrol dan dicocokkan dengan pertanian organik untuk karakteristik	Raynolds <i>et al.</i> , (2004) meneliti 'Ekspansi cepat yang sedang berlangsung dalam jaringan kopi Global Fairtrade yang menghubungkan konsumen Utara dengan petani di Selatan.' Mereka mengambil tujuh koperasi Amerika Latin (yaitu, lima di Meksiko, satu di Guatemala dan satu di El Salvador) sebagai sampel untuk menemukan faktor-faktor yang memfasilitasi keberhasilan keterlibatan koperasi dalam jaringan Fairtrade. Menurut Raynolds <i>et al.</i> , (2004) petani yang berorganisasi memperoleh keuntungan yang lebih baik dalam jaringan Fairtrade. Dalam jangka pendek, keuntungan

<p>'biofisik dan sosial ekonomi'. Karena kedekatannya dengan kelompok intervensi, pertanian konvensional dapat memberikan evaluasi yang adil terhadap praktik organik. Oleh karena itu, penilaian dampak sertifikasi organik pada petani kopi bisa lebih andal. Studi ini menemukan bahwa produksi lima pertanian organik lebih tinggi dari produksi pertanian konvensional; namun, hasil tiga pertanian organik lainnya adalah 22% lebih rendah dari pertanian konvensional. Selain itu, Lyngbaek dan Muschler (2001) berpendapat bahwa biaya sertifikasi organik akan meningkatkan biaya pertanian organik. Biaya sertifikasi membuat harga premium organik tidak dapat menutupi total biaya yang dikeluarkan oleh petani, sehingga total biaya pertanian organik jauh lebih tinggi daripada biaya pertanian konvensional. Oleh karena itu, pendapatan bersih petani organik sebenarnya lebih rendah dari pendapatan petani konvensional.</p>	<p>ekonomi akan tampak rendah; tetapi dalam jangka panjang, pemberdayaan dan peningkatan kapasitas akan mendapat manfaat dari <i>Fairtrade</i>. Namun, dalam studi ini, para peneliti tampaknya mengabaikan peran kelompok kontrol sebagai faktor kontra faktual dalam analisis mereka. Mereka juga tidak memiliki metode yang andal untuk mengevaluasi dampak jaringan <i>Fairtrade</i>. Oleh karena itu, hasil penelitian tidak bisa dikatakan bebas dari bias karena dampak yang ditimbulkan pada koperasi kopi bisa berasal dari penyebab lain, selain <i>Fairtrade</i>.</p>
<p>Bolwig et al., (2009) meneliti 'Ekonomi pertanian kontrak organik petani kecil di Afrika.' Penulis mensurvei total 160 petani di Uganda yang berpartisipasi dalam skema pertanian kontrak kopi organik Sipi milik Kawacom (U) Ltd. Mereka menemukan bahwa, dengan beberapa faktor yang terkendali, partisipasi dalam pertanian kontrak memberikan pendapatan yang lebih optimis bagi petani daripada tanpa kontrak. Partisipasi dalam skema ini dapat meningkatkan pendapatan bersih petani rata-rata 75%, sedangkan tanpa kontrak hanya meningkatkan pendapatan petani rata-rata sekitar 9%. Dalam penelitian ini, pemilihan desain baik kelompok intervensi maupun kelompok kontrol didasarkan pada metode <i>two-stage random sampling</i>. Sementara kelompok intervensi dipilih secara acak dari daftar petani yang tercatat yang diberikan oleh Kawacom, kelompok kontrol dipilih secara acak dari daftar petani yang disediakan oleh kepala desa. Kondisi agroekologi kedua kelompok (hampir) mirip satu sama lain. Selain itu, Bolwig et al., (2009) menggunakan "Regresi OLS standar dan perkiraan kemungkinan maksimum informasi lengkap (FIML) dari model seleksi Heckman dalam memberikan karakteristik yang cocok untuk membandingkan antara kelompok intervensi dan kelompok kontrol. Model ini kuat, terutama dalam menangani ukuran sampel yang kecil. Oleh karena itu, secara keseluruhan, Bolwig et al., (2009) menggunakan metode yang bertujuan untuk mengurangi bias yang sering terjadi akibat teknik yang tidak tepat dalam memilih sampel dan membandingkan objek.</p>	<p>Kilian et al., (2004) meneliti dampak dari empat proses sertifikasi (yaitu, Organik, <i>Fairtrade</i>, Rainforest Alliance, dan Utz) pada sektor ekonomi mikro dengan mempertimbangkan potensi segmen swasta dan sudut pandang pasar jangka panjang. Sementara data yang digunakan untuk analisis pasar didasarkan pada survei yang dilakukan oleh CIMS Foundation, sedangkan penilaian ekonomi didasarkan pada dua master tesis milik mahasiswa University Of Costa Rica. Kilian et al., (2004) menemukan bahwa sertifikasi memberi petani keuntungan finansial (misalnya, harga yang lebih baik) dan prospek kewirausahaan (misalnya, dalam membedakan produk mereka di pasar yang kompetitif). Kilian et al., (2004) juga menemukan bahwa, meskipun konsep mereka tampak tumpang tindih, sertifikasi-sertifikasi tersebut memiliki pendekatan yang berbeda dalam mendorong produksi kopi berkelanjutan. Menurut Kilian et al., (2004), <i>Fairtrade</i> lebih dapat diterapkan pada petani kecil karena menawarkan harga premium; namun, hal itu menuntut praktik lingkungan dan sosial yang tinggi kepada para petani. Sebaliknya, Rainforest Alliance dan Utz Kapeh lebih cocok untuk petani kopi skala besar. Kilian et al., (2004) menyatakan bahwa pasar di Amerika Utara dan Eropa hampir sama besarnya untuk kopi bersertifikat Organik. Pasar Eropa menjadi tujuan utama kopi <i>Fairtrade</i>, selain Amerika Serikat. Pasar Eropa dan Amerika Serikat masing-masing merupakan pasar utama bagi Utz Kapeh dan Rainforest Alliance. Namun demikian, metode penelitian ini tampak meragukan karena sebagian besar menggunakan data sekunder (yaitu, survei CIMS dan tesis master), penelitian ini tidak dapat mengevaluasi secara tepat</p>

	<p>pengaruh sertifikasi dengan membandingkan secara proporsional petani sertifikasi dengan petani non-sertifikasi di lapangan. Peneliti juga tidak menjelaskan apakah data sekunder yang digunakan memiliki karakteristik yang mendukung perbandingan proporsional antara kelompok intervensi dan kelompok kontrol.</p>
<p>Ruben dan Zuniga (2011) mengevaluasi dampak berbagai jenis standar terhadap kesejahteraan petani, dan peran standar ini dalam peningkatan rantai nilai. Mereka menyurvei 315 petani kopi di Nikaragua Utara. Mereka membandingkan dampak sertifikasi <i>Fairtrade</i>, Rainforest Alliance, dan Café Practices di sektor kopi. Mereka menemukan bahwa petani di bawah <i>Fairtrade</i> menerima harga yang lebih tinggi daripada petani mandiri, tetapi Café Practices tetap lebih baik daripada <i>Fairtrade</i> dalam hal produktivitas dan kualitas. Ruben dan Zuniga (2011) memberikan dua kategori kelompok kontrol untuk membandingkan pengaruh <i>Fairtrade</i> dan skema sertifikasi lain yang berbeda (yaitu, Rainforest Alliance dan Café Practices) terhadap kesejahteraan petani. Kelompok kontrol pertama adalah petani yang membudidayakan kopi di bawah Rainforest Alliance and Café Practices, dan kelompok kedua adalah petani yang melakukan cara konvensional. Mereka menggunakan sampel yang seimbang untuk setiap kelompok intervensi (petani sertifikasi) dan kelompok kontrol (petani non-sertifikasi). Teknik Propensity Score Matching (PSM) digunakan untuk menghindari bias pemilihan sampel. Oleh karena itu, metode penelitian ini kuat karena mampu mengoreksi bias seleksi sambil menilai dan membandingkan dampak dari berbagai sertifikasi.</p>	<p>Bacon (2005) mempelajari 'Dampak penjualan pada pasar organik dan <i>Fairtrade</i>' dengan menyurvei 228 petani dari lokasi sosial dan lingkungan yang berbeda, Bacon menemukan bahwa sertifikasi organik dan <i>Fairtrade</i> berdampak pada mata pencaharian petani. Peneliti menyimpulkan bahwa petani menerima harga yang lebih tinggi dan kepemilikan lahan yang lebih terjamin ketika mereka berpartisipasi dalam sertifikasi. Penelitian Bacon memiliki data yang kuat karena peneliti menyelesaikan penelitian ekstensif dengan mengumpulkan informasi dari 228 responden petani. Bacon (2005) juga mencoba memasukkan dampak dari skema sertifikasi yang berbeda - organik dan <i>Fairtrade</i> - dalam studinya. Namun, kelemahan studi Bacon adalah ia tidak memasukkan kelompok petani non-sertifikasi yang signifikan dan sebanding sebagai kelompok kontrol dalam mengevaluasi dampak sertifikasi.</p>
<p>Blackman dan Naranjo (2012) mengevaluasi 'Dampak lingkungan dari sertifikasi kopi organik di pusat Kosta Rika.' Para peneliti menemukan bahwa sertifikasi kopi organik dapat mengembangkan kinerja ekologi petani kopi dengan mendorong beberapa perubahan dalam praktik pertanian. Sertifikasi mendorong petani untuk mengurangi bahan kimia secara signifikan untuk input pertanian mereka; sementara pada saat yang sama, petani termotivasi untuk mengintensifkan penggunaan beberapa 'praktik pengelolaan yang ramah lingkungan.' Untuk studi mereka, Blackman dan Naranjo (2012) mengumpulkan sampel dari 36 pertanian organik bersertifikat dan 2.567 pertanian non-sertifikasi. Data berasal dari tiga sumber: sensus nasional petani kopi Kosta Rika yang disediakan oleh Institut Statistik dan Sensus Nasional (Instituto Nacional de Estadística y Censos,</p>	<p>Philpott et al., (2007) melakukan penelitian di Chiapas, Meksiko, di mana mereka mengevaluasi dampak ekonomi dan ekologi dari sertifikasi kopi di daerah tersebut. Mereka membandingkan delapan koperasi bersertifikat organik (organik dan <i>Fairtrade</i> organik) dengan koperasi non-sertifikasi. Para peneliti menemukan bahwa sertifikasi tidak berdampak pada penampilan vegetasi dan fauna. Namun, petani dengan sertifikasi (yaitu, organik, dan <i>Fairtrade</i> organik), dalam beberapa kasus, menerima penghasilan yang lebih tinggi daripada petani non-sertifikasi. Selain itu, Philpott et al., (2007) menyimpulkan bahwa sertifikasi tunggal tidak akan membawa dampak sertifikasi yang sukses terhadap lingkungan. Kombinasi sertifikasi organik, <i>Fairtrade</i> dan naungan akan menghasilkan pendekatan konservasi yang lebih baik di area penanaman kopi. Namun demikian, studi oleh Philpott et al., (2007) menunjukkan bahwa para</p>

<p>INEC) bekerja sama dengan Institut Kopi Kosta Rika (Instituto del Café de Costa Rica, ICAFE), data GIS disusun dari berbagai sumber, dan daftar data petani Asosiasi Petani Organik Turrialba (Asociacion de Productores Organicos de Turrialba, APOT). Untuk mengatasi 'bias seleksi diri', Blackman dan Naranjo (2012) menggunakan teknik Propensity Score Matching (PSM) dalam analisis mereka. Teknik ini berguna untuk membangun karakteristik pengamatan yang sebanding antara petani non-sertifikasi dan petani sertifikasi. Selain itu, dengan memastikan kelompok kontrol memiliki fitur dan keadaan yang serupa dengan kelompok intervensi, Blackman dan Naranjo (2012) mengukur dampak sertifikasi sebagai 'efek intervensi rata-rata pada intervensi (ATT).' Ini berarti bahwa yang disebut dampak adalah perbedaan antara persentase petani sertifikasi dan persentase petani non-sertifikasi yang menggunakan praktik pengelolaan yang sama (yaitu pertanian organik).</p>	<p>peneliti tidak melakukan uji coba terhadap ketiga kategori koperasi tani tersebut. Studi mereka kurang kelompok kontrol yang andal sebagai pesaing untuk membandingkan apakah sertifikasi berdampak pada petani. Para peneliti juga gagal untuk memisahkan secara jelas pengaruh masing-masing sertifikasi terhadap petani karena beberapa sertifikasi memiliki tujuan yang tumpang tindih. Oleh karena itu, dampak independen dan simultan dari satu sertifikasi sulit diukur dan kemungkinan besar akan bias dalam hasilnya.</p>
<p>Ruben dan Fort (2012) mempelajari 'Dampak sertifikasi <i>Fairtrade</i> bagi petani kopi di Peru.' Mereka menemukan bahwa tidak ada perbedaan yang signifikan, dalam hal harga riil, antara petani sertifikasi <i>Fairtrade</i> dan petani konvensional. Dengan demikian, sertifikasi hanya memberikan 'pendapatan langsung yang terbatas dan efek produksi' kepada para petani. Namun, sertifikasi mendorong perubahan substansial dalam beberapa aspek di luar kesejahteraan petani seperti peningkatan koperasi, penggunaan input, permodalan dan kepemilikan, dan sikap risiko. Ruben dan Fort (2012) menggunakan teknik Propensity Score Matching (PSM) dalam penelitian mereka untuk mengidentifikasi kelompok petani konvensional yang sebanding, dalam semua atribut pra-intervensi, dengan petani sertifikasi. Oleh karena itu, metode penelitian ini andal karena menggunakan perbandingan yang memadai antara petani sertifikasi <i>Fairtrade</i> (yaitu, sebagai kelompok intervensi) dan petani yang non-sertifikasi (yaitu, sebagai kelompok kontrol).</p>	<p>Bacon et al. ,(2008) melakukan studi di Nikaragua dengan menyurvei 177 rumah tangga petani yang menjalin perdagangan kopi dengan pasar konvensional dan <i>Fairtrade</i>. 177 rumah tangga terdiri dari 101 anggota serikat koperasi CECOCAFEN yang mengikuti sertifikasi <i>Fairtrade</i>, 61 petani yang terhubung ke pasar konvensional, dan 15 petani kopi organik bersertifikat. Penelitian ini bertujuan untuk menganalisis dampak <i>Fairtrade</i> terhadap kehidupan petani dengan membandingkan sosial ekonomi kehidupan petani dengan kerangka Millennium Development Goals (MDGs). Peneliti menemukan bahwa Koperasi <i>Fairtrade</i> memberikan dampak positif bagi rumah tangga dalam hal pendidikan, aset infrastruktur, dan investasi moneter. Namun, Bacon et al. ,(2008) tidak memfokuskan evaluasi mereka pada dampak yang jelas dari <i>Fairtrade</i> terhadap kehidupan Petani. Hal ini dikarenakan peneliti hanya melakukan perbandingan berdasarkan pasar di mana petani menjual kopinya. Oleh karena itu, penelitian Bacon et al. ,(2008) kehilangan penjelasan tentang karakteristik penting dari kelompok kontrol yang cocok di mana kelompok kontrol harus memiliki karakteristik yang sebanding dengan kelompok intervensi. Namun, kelompok kontrol melakukan praktik yang berbeda dari kelompok intervensi. Selain itu, dalam hal ini Bacon et al. ,(2008) membandingkan rumah tangga, anggota koperasi yang dijual ke Pasar <i>Fairtrade</i>, dengan yang lain yang dijual ke pasar konvensional. Cara ini kemungkinan akan menimbulkan pertanyaan apakah kedua rumah tangga tersebut secara konsisten menjual kopi mereka</p>

<p>Van Rijsbergen et al., (2016) mengevaluasi dampak sertifikasi kopi Utz dan <i>Fairtrade</i> di Kenya. Mereka menemukan bahwa kesejahteraan rumah tangga dan dampak sertifikasi kopi pada mata pencaharian petani umumnya tetap mengecewakan. Petani sertifikasi Utz pada awalnya menerima harga yang jauh lebih tinggi untuk buah kopi basah dan biji kopi kering, tetapi perbedaan harga ini menghilang seiring waktu karena semua kopi bersertifikat meningkat harganya. Petani sertifikasi mampu memperbaiki situasi mereka terutama melalui harga yang lebih baik untuk kopi olahan (kering), dengan penurunan perbedaan antara sertifikasi Utz dan FT tetapi meningkatkan margin dibandingkan dengan petani non-sertifikasi. Namun, petani non-sertifikasi juga dapat menyesuaikan praktik pertanian kopi mereka dari waktu ke waktu dan menghasilkan pendapatan yang cukup besar dari kegiatan pertanian dan non-pertanian lainnya. Oleh karena itu, pengaruh sertifikasi terhadap pendapatan rumah tangga masih cukup terbatas, karena rata-rata tidak lebih dari sepertiga hingga seperempat pendapatan pertanian diperoleh dari penjualan kopi. Efek pendapatan bersih yang diharapkan dari sertifikasi kurang dari 10% dari pendapatan rumah tangga petani. Dari segi metode, penelitian Van Rijsbergen et al., (2016) ini melakukan kerja lapangan dengan dua putaran pengumpulan data, mengunjungi petani yang sama pada tahun 2009 dan 2013. Petani sertifikasi dari koperasi bersertifikat (kelompok intervensi) dan dibandingkan dengan petani non-sertifikasi dari koperasi non-bersertifikat (kelompok kontrol). Koperasi non-bersertifikat yang dipilih memiliki karakteristik yang sama dengan koperasi bersertifikat dalam hal jumlah pabrik pengolahan, keanggotaan, dan struktur tata kelolanya.</p>	<p>ke pasar yang sama, atau terkadang mereka berpindah pasar.</p> <p>Valkila (2009) mengevaluasi 'Dampak produksi kopi organik <i>Fairtrade</i> terhadap kesejahteraan petani skala kecil di Nikaragua'. Peneliti melakukan wawancara semi struktur kepada sejumlah petani, perwakilan koperasi, perusahaan ekspor, pemerintah dan LSM, dan lembaga sertifikasi selama tujuh bulan kerja lapangan antara tahun 2005 dan 2008. Valkila (2009) menemukan bahwa produksi <i>Fairtrade</i> organik meningkatkan petani. Namun, karena produksi pertanian organik rendah, keuntungan ekonomi dari sertifikasi <i>Fairtrade</i> organik 'sangat terbatas'. Selain itu, harga di pasar secara signifikan menentukan manfaat finansial dari sertifikasi. Dari segi metode, penelitian ini memiliki keterbatasan karena jumlah petani yang non-sertifikasi yang bertindak sebagai kelompok kontrol terbatas; sebaliknya bukti mengenai praktik pertanian konvensional dikumpulkan dari beberapa sumber lain. Karena rendahnya produksi organik dan ketidakpastian harga pasar, penelitian Valkila (2009) ini juga gagal memberikan kesimpulan yang kuat apakah sertifikasi <i>Fairtrade</i> organik memiliki dampak positif yang signifikan terhadap kehidupan Petani.</p>
---	---

Sumber: hasil kajian literatur

Kesimpulan

Penelitian ini memiliki tiga tujuan. Pertama, meninjau kembali definisi evaluasi dampak menurut literatur. Kedua, mengidentifikasi sumber bias dalam evaluasi dampak. Ketiga, meninjau kembali studi-studi tentang evaluasi dampak, yang memiliki hasil berbeda, terutama dari metodologi yang digunakan.

Salah satu definisi evaluasi dampak yang paling umum adalah evaluasi yang berkaitan dengan penetapan kontra faktual, yaitu perbedaan yang dibuat oleh suatu program intervensi dibandingkan dengan bila program intervensi tersebut tidak ada. Kontra faktual yang tidak didesain dengan tepat merupakan sumber utama bias dalam evaluasi dampak, selain ancaman validitas. Berdasarkan kajian studi-studi empiris tentang dampak standar dan sertifikasi kopi, disimpulkan bahwa sebagian studi hasilnya lebih dapat dipercaya dibandingkan studi-studi yang lain. Studi-studi yang hasilnya lebih dapat dipercaya mempergunakan desain kontra faktual yang

andal (*reliable*), sehingga mampu meminimalkan bias evaluasi dampak melalui perbandingan yang sesuai (*apple-to-apple comparison*) antara kelompok intervensi dan kelompok kontrol.

Dengan demikian, penelitian ini bermanfaat terutama bagi para peneliti lainnya. Para peneliti bertujuan mencari kebenaran, dan metode yang andal dan dapat meminimalkan bias merupakan penunjang untuk mencapai tujuan itu.

Daftar Pustaka

- Bacon C. 2005. Confronting the Coffee Crisis: Can Fair Trade, Organic, and Specialty Coffees Reduce Small-Scale Farmer Vulnerability in Northern Nicaragua? *World Dev.* 33(3):497–511.
- Bacon CM, Ernesto Mendez V, Gómez MEF, Stuart D, Flores SRD. 2008. Are sustainable coffee certifications enough to secure farmer livelihoods? The millenium development goals and Nicaragua's Fair Trade cooperatives. *Globalizations.* 5(2):259–274.
- Blackman A, Naranjo MA. 2012. Does eco-certification have environmental benefits? Organic coffee in Costa Rica. *Ecol. Econ.* 83:58–66.
- Blackman A, Rivera J. 2011. Producer-Level Benefits of Sustainability Certification. *Conserv. Biol.* 25(6):1176–1185.doi:10.1111/j.1523-1739.2011.01774.x.
- Bolwig S, Gibbon P, Jones S. 2009. The economics of smallholder organic contract farming in tropical Africa. *World Dev.* 37(6):1094–1104.
- Calo M, Wise TA. 2005. Revaluing peasant coffee production: Organic and fair trade markets in Mexico. *Glob. Dev. Environ. Institute, Tufts Univ.*
- Crato N, Paruolo P. 2018. *Data-driven policy impact evaluation: How access to microdata is transforming policy design.* Springer Nature.
- Dietz T. 2021. Social impact assessment as a tool for rangeland management. Di dalam: *Developing Strategies for Rangeland Management.* CRC Press. hlm. 1613–1634.
- Dietz T, Estrella Chong A, Grabs J, Kilian B. 2020. How effective is multiple certification in improving the economic conditions of smallholder farmers? Evidence from an impact evaluation in Colombia's Coffee Belt. *J. Dev. Stud.* 56(6):1141–1160.
- Ferraro PJ. 2009. Counterfactual thinking and impact evaluation in environmental policy. *New Dir. Eval.* 2009(122):75–84.doi:10.1002/ev.297.
- Gertler PJ, Martinez S, Premand P, Rawlings LB, Vermeersch CMJ. 2016. *Impact evaluation in practice.* World Bank Publications.
- Glasbergen P. 2018. Smallholders do not eat certificates. *Ecol. Econ.* 147:243–252.
- Grabs J. 2020. Assessing the institutionalization of private sustainability governance in a changing coffee sector. *Regul. Gov.* 14(2):362–387.
- International Initiative for Impact Evaluation (3ie). 2008. *Principles for Impact Evaluation.* 3ie New Delhi.
- Khandker S, B. Koolwal G, Samad H. 2009. *Handbook on Impact Evaluation.* Washington DC: The World Bank.
- Kilian B, Pratt L, Jones C, Villalobos A. 2004. Can the private sector be competitive and contribute to development through sustainable agricultural business? A case study of coffee in Latin America. *Int. Food Agribus. Manag. Rev.* 7(1030-2016–82554):21–45.
- Lyngbæk AE, Muschler RG. 2001. Productivity and profitability of multistrata organic versus conventional coffee farms in Costa Rica. *Agrofor. Syst.* 53(2):205–213.
- Maclellan N. 2016. Regional Trade Agreements in the Pacific Islands: Fair Trade for Farmers? Di dalam: *Fair Trade, Corporate Accountability and Beyond.* Routledge. hlm. 313–332.
- Mancini L, Sala S. 2018. Social impact assessment in the mining sector: Review and comparison of indicators frameworks. *Resour. Policy.* 57:98–111.
- McGill E, Er V, Penney T, Egan M, White M, Meier P, Whitehead M, Lock K, de Cuevas RA, Smith R. 2021. Evaluation of public health interventions from a complex systems perspective: a research methods review. *Soc. Sci. Med.* 272:113697.
- Meemken E-M, Barrett CB, Michelson HC, Qaim M, Reardon T, Sellare J. 2021. Sustainability standards in global agrifood supply chains. *Nat. Food.* 2(10):758–765.
- Philpott SM, Bichier P, Rice R, Greenberg R. 2007. Field-testing ecological and economic benefits of coffee certification programs. *Conserv. Biol.* 21(4):975–985.
- Ranjan Jena P, Grote U. 2017. Fairtrade certification and livelihood impacts on small-scale coffee producers in a tribal community of India. *Appl. Econ. Perspect. Policy.* 39(1):87–110.
- Raynolds LT. 2004. The Globalization of Organic Agro-Food Networks. *World Dev.* 32(5):725–743.
- Raynolds LT, Murray D, Leigh Taylor P. 2004. Fair trade coffee: building producer capacity via global networks. *J. Int. Dev. J. Dev. Stud. Assoc.* 16(8):1109–1121.
- Riedl E. 2009. Fair trade: Evaluating opportunities for smallholder coffee growers in Papua New Guinea. *Pacific Econ. Bull.* 24(2):122–140.

- Van Rijsbergen B, Elbers W, Ruben R, Njuguna SN. 2016. The ambivalent impact of coffee certification on farmers' welfare: a matched panel approach for cooperatives in Central Kenya. *World Dev.* 77:277–292.
- Ruben R, Fort R. 2012. The Impact of Fair Trade Certification for Coffee Farmers in Peru. *World Dev.* 40(3):570–582.doi:10.1016/j.worlddev.2011.07.030.
- Ruben R, Zuniga G. 2011. How standards compete: comparative impact of coffee certification schemes in Northern Nicaragua. *Supply Chain Manag. An Int. J.* 16(2):98–109.doi:10.1108/13598541111115356.
- https://www.elgaronline.com/abstract/edcoll/9781782549550/9781782549550.00020.xmlSchre%20L.%202016.%20Social%20impact%20assessment.%20Handb.%20Regul.%20Impact%20Assess.:108–126.doi:10.4337/9781782549567.00020.%20%5bdiunduh%202003%20Mei%2022%5d.%20Tersedia%20pada:%20http://www.gsvc-sea.org/?page=social_impact_assessment
- Schweizer ML, Braun BI, Milstone AM. 2016. Research methods in healthcare epidemiology and antimicrobial stewardship—quasi-experimental designs. *Infect. Control Hosp. Epidemiol.* 37(10):1135–1140.
- the Organization for Economic Co-operation and Development. 2022. Outline Of Principles Of Impact Evaluation. [diunduh 2022 Jun 15]. Tersedia pada: <https://www.oecd.org/dac/evaluation/dcdndep/37671602.pdf>
- Traldi R. 2021. Progress and pitfalls: A systematic review of the evidence for agricultural sustainability standards. *Ecol. Indic.* 125:107490.
- US Environmental Protection Agency. 2022. Program Evaluation Glossary. [diunduh 2022 Jun 25]. Tersedia pada: [https://sor.epa.gov/sor_internet/registry/termreg/searchandretrieve/glossariesandkeywordlists/search.do?details=&glossaryName=Program Evaluation Glossary](https://sor.epa.gov/sor_internet/registry/termreg/searchandretrieve/glossariesandkeywordlists/search.do?details=&glossaryName=Program%20Evaluation%20Glossary)
- Valkila J. 2009. Fair Trade organic coffee production in Nicaragua—Sustainable development or a poverty trap? *Ecol. Econ.* 68(12):3018–3025.
- White H. 2006. *Impact evaluation: the experience of the Independent Evaluation Group of the World Bank*. Munich, Germany: Independent Evaluation Group, the World Bank.
- Williams CC, Kedir AM. 2020. Evaluating the impact of registration on future firm performance in the Middle East and North Africa region: evidence from the World Bank Enterprise Survey. *Int. J. Entrep. Small Bus.* 41(4):473–489.
- Williams MJ. 2020. External validity and policy adaptation: From impact evaluation to policy design. *World Bank Res. Obs.* 35(2):158–191.