Online appendix K

Energy sector risk of bias assessment report

Table 1. Risk of bias assessment in the energy sector

ITEM	1A. STUDY DESIGN, END OF INTERVENTION (POTENTIAL CONFOUNDERS TAKEN INTO ACCOUNT)	4. 1NTERVENTION IS CLEARLY DEFINED	5. OUTCOME MEASURES ARE CLEARLY DEFINED AND RELIABLE	6. BASELINE BALANCE (N.A. FOR BEFORE VERSUS AFTER)	7. REPRESENTATIVENESS OF A LARGE SCALE INTERVENTION	7A. IS THE SAMPLE FRAME CLEARLY DEFINED?	8. PRECISION OF ESTIMATE (IN CASE OF REGRESSION)	OVERALL CONFIDENCE IN STUDY (END OF INTERVENTION)	STUDY QUALITY
(TANG 2020) (ID:56954267)	-Medium confidence: DiD with matching, PSM	-High confidence: Intervention clearly and fully described	-High confidence: Outcome measure clearly and fully described, preferably with reference to validation	-Medium confidence: Imbalance in 5-10 percent baseline variables	-Medium confidence: At least 7b is answered with a "YES" and 7c is not answered as a "NO"	-Yes	-N/A for studies where no clusters are mentioned	-Medium confidence: Medium rating across any of the items 1a, 4, 5, 6, 7 and 8 (and NO low rating)	-Medium Quality Impact Evaluation
AKPANDJAR (2017) (ID:47518460)	-Low confidence: Other matching, DID alone,	-Low confidence: Intervention named but not	-High confidence: Outcome measure clearly and fully described, preferably with	-High confidence: RCT or baseline balance report and satisfactory	-High confidence: All three (7a,7b & 7c) answered as a "YES"	-Yes	-High confidence: Standard errors are clustered by	-Low confidence: Low rating across any of the items 1a, 4, 5, 6, 7 and 8	-Low Quality Impact Evaluation

		described, or not named	reference to validation	(imbalance on 5 or less than 5 percent)			intervention design level		
instrumental variable otherwise									
CHAKRAVORTY (2016) (ID:48097585)	-Medium confidence: DiD with matching, PSM	-High confidence: Intervention clearly and fully described	-Medium confidence: Outcome named but not described	-Medium confidence: Imbalance in 5-10 percent baseline variables	-Low confidence: None is a "YES" (7a,7b,7c)	-Yes	-High confidence: Standard errors are clustered by intervention design level	-Medium confidence: Medium rating across any of the items 1a, 4, 5, 6, 7 and 8 (and NO low rating)	-Medium Quality Impact Evaluation
DASSO (2015) (ID:48145174)	-High confidence: RCT, RDD, instrumental variable under LATE	-High confidence: Intervention clearly and fully described	-High confidence: Outcome measure clearly and fully described, preferably with reference to validation	-High confidence: RCT or baseline balance report and satisfactory (imbalance on 5 or less than 5 percent)	-High confidence: All three (7a,7b & 7c) answered as a "YES"	-Yes	-High confidence: Standard errors are clustered by intervention design level	-High confidence: High rating across ALL of the items 1a, 4, 5, 6, 7 and 8 (or N/A in 8)	-High Quality Impact Evaluation
DINKELMAN (2011) (ID:48093940)	-Low confidence: Other matching, DID alone, instrumental variable otherwise	-High confidence: Intervention clearly and fully described	-High confidence: Outcome measure clearly and fully described, preferably with reference to validation	-Low confidence: Baseline balance not reported or reported and lack of balance on 10 Or more than 10% of baseline variables (except RCT)	-Medium confidence: At least 7b is answered with a "YES" and 7c is not answered as a "NO"	-Yes	-High confidence: Standard errors are clustered by intervention design level	-Low confidence: Low rating across any of the items 1a, 4, 5, 6, 7 and 8	-Low Quality Impact Evaluation
DONG (2019) (ID:56954272)	-High confidence: RCT, RDD, instrumental variable under LATE	-High confidence: Intervention clearly and fully described	-High confidence: Outcome measure clearly and fully described, preferably with reference to validation	-High confidence: RCT or baseline balance report and satisfactory (imbalance on 5 or less than 5 percent)	-Medium confidence: At least 7b is answered with a "YES" and 7c is not answered as a "NO"	-Yes	-High confidence: Standard errors are clustered by intervention design level	-Medium confidence: Medium rating across any of the items 1a, 4, 5, 6, 7 and 8 (and NO low rating)	-Medium Quality Impact Evaluation
DONG (2020) (ID:56954264)	-Medium confidence: DiD with matching, PSM	-High confidence: Intervention clearly and fully described	-High confidence: Outcome measure clearly and fully described, preferably with reference to validation	-Low confidence: Baseline balance not reported or reported and lack of balance on 10 Or more than 10% of baseline variables (except RCT)	-High confidence: All three (7a,7b & 7c) answered as a "YES"	-Yes	-N/A for studies where no clusters are mentioned	-Low confidence: Low rating across any of the items 1a, 4, 5, 6, 7 and 8	-Low Quality Impact Evaluation
GAO (2020) (ID:56954265)	-Medium confidence:	-High confidence:	-High confidence: Outcome measure	-Medium confidence: Imbalance in 5-10	-Medium confidence: At least 7b is	-Yes	-N/A for studies	-Medium confidence:	-Medium Quality

	DiD with matching, PSM	Intervention clearly and fully described	clearly and fully described, preferably with reference to validation	percent baseline variables	answered with a "YES" and 7c is not answered as a "NO"		where no clusters are mentioned	Medium rating across any of the items 1a, 4, 5, 6, 7 and 8 (and NO low rating)	Impact Evaluation
GROGAN (2013) (ID:47518519)	-Medium confidence: DiD with matching, PSM	-High confidence: Intervention clearly and fully described	-High confidence: Outcome measure clearly and fully described, preferably with reference to validation	-Medium confidence: Imbalance in 5-10 percent baseline variables	-Medium confidence: At least 7b is answered with a "YES" and 7c is not answered as a "NO"	-Yes	-N/A for studies where no clusters are mentioned	-Medium confidence: Medium rating across any of the items 1a, 4, 5, 6, 7 and NO low rating)	-Medium Quality Impact Evaluation
GROGAN (2016) (ID:48100131)	-Low confidence: Other matching, DID alone, instrumental variable otherwise	-High confidence: Intervention clearly and fully described	-High confidence: Outcome measure clearly and fully described, preferably with reference to validation	-Low confidence: Baseline balance not reported or reported and lack of balance on 10 0r more than 10% of baseline variables (except RCT)	-High confidence: All three (7a,7b & 7c) answered as a "YES"	-Yes	-High confidence: Standard errors are clustered by intervention design level	-Low confidence: Low rating across any of the items 1a, 4, 5, 6, 7 and 8	-Low Quality Impact Evaluation
GROGAN (2018) (ID:48097895)	-High confidence: RCT, RDD, instrumental variable under LATE	-High confidence: Intervention clearly and fully described	-High confidence: Outcome measure clearly and fully described, preferably with reference to validation	-Low confidence: Baseline balance not reported or reported and lack of balance on 10 0r more than 10% of baseline variables (except RCT)	-Medium confidence: At least 7b is answered with a "YES" and 7c is not answered as a "NO"	-Yes	-High confidence: Standard errors are clustered by intervention design level	-Low confidence: Low rating across any of the items 1a, 4, 5, 6, 7 and 8	-Low Quality Impact Evaluation
HU (2020) (ID:48094926)	-High confidence: RCT, RDD, instrumental variable under LATE	-High confidence: Intervention clearly and fully described	-High confidence: Outcome measure clearly and fully described, preferably with reference to validation	-Low confidence: Baseline balance not reported or reported and lack of balance on 10 0r more than 10% of baseline variables (except RCT)	-Medium confidence: At least 7b is answered with a "YES" and 7c is not answered as a "NO"	-Yes	-High confidence: Standard errors are clustered by intervention design level	-Low confidence: Low rating across any of the items 1a, 4, 5, 6, 7 and 8	-Low Quality Impact Evaluation
LEE (2020) (ID:48100451)	-High confidence: RCT, RDD, instrumental variable under LATE	-High confidence: Intervention clearly and fully described	-High confidence: Outcome measure clearly and fully described, preferably with reference to validation	-High confidence: RCT or baseline balance report and satisfactory (imbalance on 5 or less than 5 percent)	-Medium confidence: At least 7b is answered with a "YES" and 7c is not answered as a "NO"	-Yes	-High confidence: Standard errors are clustered by intervention design level	-Medium confidence: Medium rating across any of the items 1a, 4, 5, 6, 7 and 8 (and NO low rating)	-Medium Quality Impact Evaluation

Study									
LENZ (2017) (ID:48100456)	-High confidence: RCT, RDD, instrumental variable under LATE	-High confidence: Intervention clearly and fully described	-High confidence: Outcome measure clearly and fully described, preferably with reference to validation	-High confidence: RCT or baseline balance report and satisfactory (imbalance on 5 or less than 5 percent)	-High confidence: All three (7a,7b & 7c) answered as a "YES"	-Yes	-High confidence: Standard errors are clustered by intervention design level	-High confidence: High rating across ALL of the items 1a, 4, 5, 6, 7 and 8 (or N/A in 8)	-High Quality Impact Evaluation
LIPSCOMB (2013) (ID:47518555)	-Medium confidence: DiD with matching, PSM	-High confidence: Intervention clearly and fully described	-High confidence: Outcome measure clearly and fully described, preferably with reference to validation	-High confidence: RCT or baseline balance report and satisfactory (imbalance on 5 or less than 5 percent)	-High confidence: All three (7a,7b & 7c) answered as a "YES"	-Yes	-High confidence: Standard errors are clustered by intervention design level	-Medium confidence: Medium rating across any of the items 1a, 4, 5, 6, 7 and 8 (and NO low rating)	-Medium Quality Impact Evaluation
LITZOW (2019) (ID:48098270)	-High confidence: RCT, RDD, instrumental variable under LATE	-High confidence: Intervention clearly and fully described	-High confidence: Outcome measure clearly and fully described, preferably with reference to validation	-Low confidence: Baseline balance not reported or reported and lack of balance on 10 Or more than 10% of baseline variables (except RCT)	-High confidence: All three (7a,7b & 7c) answered as a "YES"	-Yes	-High confidence: Standard errors are clustered by intervention design level	-High confidence: High rating across ALL of the items 1a, 4, 5, 6, 7 and 8 (or N/A in 8)	-High Quality Impact Evaluation
QI (2021) (ID:56954266)	-Medium confidence: DiD with matching, PSM	-High confidence: Intervention clearly and fully described	-High confidence: Outcome measure clearly and fully described, preferably with reference to validation	-High confidence: RCT or baseline balance report and satisfactory (imbalance on 5 or less than 5 percent)	-High confidence: All three (7a,7b & 7c) answered as a "YES"	-Yes	-High confidence: Standard errors are clustered by intervention design level	-Medium confidence: Medium rating across any of the items 1a, 4, 5, 6, 7 and 8 (and NO low rating)	-Medium Quality Impact Evaluation
SAMAD (2016) (ID:48145176)	-Medium confidence: DiD with matching, PSM	-High confidence: Intervention clearly and fully described	-High confidence: Outcome measure clearly and fully described, preferably with reference to validation	-Low confidence: Baseline balance not reported or reported and lack of balance on 10 Or more than 10% of baseline variables (except RCT)	-High confidence: All three (7a,7b & 7c) answered as a "YES"	-Yes	-High confidence: Standard errors are clustered by intervention design level	-Low confidence: Low rating across any of the items 1a, 4, 5, 6, 7 and 8	-Medium Quality Impact Evaluation
SAMAD (2017) (ID:48145133)	-Low confidence: Other matching, DID alone, instrumental	-High confidence: Intervention clearly and fully described	-High confidence: Outcome measure clearly and fully described, preferably with	-Low confidence: Baseline balance not reported or reported and lack of balance on 10 Or more than 10%	-High confidence: All three (7a,7b & 7c) answered as a "YES"	-Yes	-High confidence: Standard errors are clustered by	-High confidence: High rating across ALL of the items 1a, 4, 5, 6, 7 and 8 (or N/A in 8)	-High Quality Impact Evaluation

Study									
	variable otherwise		reference to validation	of baseline variables (except RCT)			intervention design level		
SAMAD (2019) (ID:48098823)	-Medium confidence: DiD with matching, PSM	-High confidence: Intervention clearly and fully described	-High confidence: Outcome measure clearly and fully described, preferably with reference to validation	-High confidence: RCT or baseline balance report and satisfactory (imbalance on 5 or less than 5 percent)	-High confidence: All three (7a,7b & 7c) answered as a "YES"	-Yes	-High confidence: Standard errors are clustered by intervention design level	-High confidence: High rating across ALL of the items 1a, 4, 5, 6, 7 and 8 (or N/A in 8)	-High Quality Impact Evaluation
SHEN (2020) (ID:56954295)	-Medium confidence: DiD with matching, PSM	-Low confidence: Intervention named but not described, or not named	-Low confidence: Brief description of outcome	-High confidence: RCT or baseline balance report and satisfactory (imbalance on 5 or less than 5 percent)	-High confidence: All three (7a,7b & 7c) answered as a "YES"	-Yes	-High confidence: Standard errors are clustered by intervention design level	-Low confidence: Low rating across any of the items 1a, 4, 5, 6, 7 and 8	-High Quality Impact Evaluation
TAGLIAPIETRA (2020) (ID:56954271)	-Low confidence: Other matching, DID alone, instrumental variable otherwise	-High confidence: Intervention clearly and fully described	-High confidence: Outcome measure clearly and fully described, preferably with reference to validation	-Low confidence: Baseline balance not reported or reported and lack of balance on 10 0r more than 10% of baseline variables (except RCT)	-High confidence: All three (7a,7b & 7c) answered as a "YES"	-Yes	-N/A for studies where no clusters are mentioned	-Low confidence: Low rating across any of the items 1a, 4, 5, 6, 7 and 8	-Low Quality Impact Evaluation
WANG (2019) (ID:48101304)	-Medium confidence: DiD with matching, PSM	-High confidence: Intervention clearly and fully described	-High confidence: Outcome measure clearly and fully described, preferably with reference to validation	-High confidence: RCT or baseline balance report and satisfactory (imbalance on 5 or less than 5 percent)	-High confidence: All three (7a,7b & 7c) answered as a "YES"	-Yes	-High confidence: Standard errors are clustered by intervention design level	-Medium confidence: Medium rating across any of the items 1a, 4, 5, 6, 7 and 8 (and NO low rating)	-Medium Quality Impact Evaluation
WANG (2020) (ID:56954273)	-Medium confidence: DiD with matching, PSM	-High confidence: Intervention clearly and fully described	-High confidence: Outcome measure clearly and fully described, preferably with reference to validation	-Low confidence: Baseline balance not reported or reported and lack of balance on 10 0r more than 10% of baseline variables (except RCT)	-High confidence: All three (7a,7b & 7c) answered as a "YES"	-Yes	-N/A for studies where no clusters are mentioned	-Low confidence: Low rating across any of the items 1a, 4, 5, 6, 7 and 8	-Low Quality Impact Evaluation
WEN (2020) (ID:56954268)	-Medium confidence:	-High confidence: Intervention	-High confidence: Outcome measure clearly and fully	-High confidence: RCT or baseline balance report and	-High confidence: All three (7a,7b & 7c) answered as a "YES"	-Yes	-N/A for studies where no	-Medium confidence: Medium rating	-Medium Quality

	DiD with matching, PSM	clearly and fully described	described, preferably with reference to validation	satisfactory (imbalance on 5 or less than 5 percent)			clusters are mentioned	across any of the items 1a, 4, 5, 6, 7 and 8 (and NO low rating)	Impact Evaluation
YI (2020) (ID:48101426)	-Medium confidence: DiD with matching, PSM	-High confidence: Intervention clearly and fully described	-High confidence: Outcome measure clearly and fully described, preferably with reference to validation	-Low confidence: Baseline balance not reported or reported and lack of balance on 10 Or more than 10% of baseline variables (except RCT)	-High confidence: All three (7a,7b & 7c) answered as a "YES"	-Yes	-N/A for studies where no clusters are mentioned	-Low confidence: Low rating across any of the items 1a, 4, 5, 6, 7 and 8	-Low Quality Impact Evaluation
ZHANG (2019) (ID:47518645)	-Low confidence: Other matching, DID alone, instrumental variable otherwise	-High confidence: Intervention clearly and fully described	-High confidence: Outcome measure clearly and fully described, preferably with reference to validation	-Low confidence: Baseline balance not reported or reported and lack of balance on 10 Or more than 10% of baseline variables (except RCT)	-High confidence: All three (7a,7b & 7c) answered as a "YES"	-Yes	-Medium confidence: Standard errors are clustered but clear at what level	-Low confidence: Low rating across any of the items 1a, 4, 5, 6, 7 and 8	-Low Quality Impact Evaluation
ZHANG (2019) (ID:48101519)	-Medium confidence: DiD with matching, PSM	-Medium confidence: Brief description of intervention	-High confidence: Outcome measure clearly and fully described, preferably with reference to validation	-Low confidence: Baseline balance not reported or reported and lack of balance on 10 Or more than 10% of baseline variables (except RCT)	-High confidence: All three (7a,7b & 7c) answered as a "YES"	-Yes	-High confidence: Standard errors are clustered by intervention design level	-Medium confidence: Medium rating across any of the items 1a, 4, 5, 6, 7 and 8 (and NO low rating)	-Medium Quality Impact Evaluation
ZHANG (2020) (ID:48094916)	-Low confidence: Other matching, DID alone, instrumental variable otherwise	-High confidence: Intervention clearly and fully described	-High confidence: Outcome measure clearly and fully described, preferably with reference to validation	-Medium confidence: Imbalance in 5-10 percent baseline variables	-High confidence: All three (7a,7b & 7c) answered as a "YES"	-Yes	-Medium confidence: Standard errors are clustered but clear at what level	-Low confidence: Low rating across any of the items 1a, 4, 5, 6, 7 and 8	-Low Quality Impact Evaluation
ZHANG (2020) (ID:56954269)	-Low confidence: Other matching, DID alone, instrumental	-High confidence: Intervention clearly and fully described	-High confidence: Outcome measure clearly and fully described, preferably with reference to validation	-Low confidence: Baseline balance not reported or reported and lack of balance on 10 Or more than 10% of baseline variables (except RCT)	-High confidence: All three (7a,7b & 7c) answered as a "YES"	-Yes	-Low confidence: Standard errors are not clustered	-Low confidence: Low rating across any of the items 1a, 4, 5, 6, 7 and 8	-Low Quality Impact Evaluation

	variable otherwise								
ZHANG (2020) (ID:56954270)	-Medium confidence: DiD with matching, PSM	-High confidence: Intervention clearly and fully described	-High confidence: Outcome measure clearly and fully described, preferably with reference to validation	-High confidence: RCT or baseline balance report and satisfactory (imbalance on 5 or less than 5 percent)	-High confidence: All three (7a,7b & 7c) answered as a "YES"	-Yes	-High confidence: Standard errors are clustered by intervention design level	-Medium confidence: Medium rating across any of the items 1a, 4, 5, 6, 7 and 8 (and NO low rating)	-Medium Quality Impact Evaluation