AUGUST 2021

CEDIL Methods Brief 3

Combining economic modelling and randomised controlled trials:

An unexploited synergy

Centre of Excellence for Development Impact and Learning

Box 1: Highlights

Economic models and experiments should be complementary, rather than alternative approaches. Both approaches have advantages and limitations but their combined use enhances policy analysis.

Randomised controlled trials are mostly silent about the impact of interventions in contexts other than that of implementation, and neither do they inform about alternative courses of action. Economic models, on the other hand, can answer an unlimited number of policy questions but are based on greatly simplified assumptions.

Economic models can be validated by testing their predictions against data from rigorous experiments. The use of economic modelling alongside randomised experiments has the ability to answer a wide range of policy questions. In particular:

- Models can show how interventions work by uncovering key behavioural mechanisms. They help answer the question of why an intervention has worked or not worked.
- Models can unpack the impact of bundled interventions. They can unpack the effectiveness of project components and the complementarities between activities.
- Models can forecast the impact of alternative courses of action. The formalisation of behavioural mechanisms behind the operation of interventions allows the simulation of interventions that were never implemented.
- Models encourage interdisciplinary research and stakeholder engagement. No fruitful model can be built if teams do not include researchers from different disciplines and policymakers.
- Models can greatly contribute to a broader understanding of the policy problem at hand by involving a diverse set of stakeholders.

Introduction

The COVID-19 pandemic has revived interest in models of social behaviour. During the pandemic, governments and the public became familiar with mathematical and epidemiological modelling of disease transmission. This modelling employs a system of equations to make predictions about the spread of infections and deaths. Few people are familiar with the mathematical formulation of the model, yet many concepts of this model have been employed by commentators, politicians, and the media to communicate with the public. People have become familiar with the concepts of a threshold 'herd immunity', 'flattening the curve', and 'reducing the R number', not realising that these are sophisticated parameters of a complex abstract model. The model has informed lockdown decisions and precautionary measures taken against the pandemic. It has had a tremendous impact on policy and it is not surprising that many social scientists have since decided to improve our understanding of the pandemic using models. This has led to novel research integrating epidemiological and economic models to weigh the trade-offs between lives and livelihoods. We may see this renewed interest in modelling sustaining in public policy and in international development.

Over the last 20 years, randomised controlled trials have been used frequently in the impact evaluation of development interventions. This development has led to a large body of rigorous empirical evidence in many areas and has supported evidence-informed policies. However, many researchers and policymakers have come to realise the limitations of randomised trials.

Randomised controlled trials produce unbiased estimates of project effects, and there is a general consensus in the scientific community that they are a useful approach to find out whether an intervention works. The results they produce are local to the context in which they are implemented, and applying the results to other environments needs to be done cautiously. According to James Heckman, randomised controlled trials are only able to respond effectively to one of three fundamental policy evaluation questions: they address the problem of internal validity by answering the question of whether an intervention has an impact on welfare or not. They are not designed to address the other two policy evaluation problems: forecasting the impacts of interventions implemented in one environment in other environments (also known as the problem of 'external validity'), or forecasting the impact of alternative, untested interventions.

To address the problem of external validity, randomised controlled trials can rely on qualitative work through process evaluations, building on existing social science theory, or multiple trials through multi-country trials and meta-analyses (as is standard in the medical sciences). However, there are limitations in terms of how far and under what circumstances this can be done robustly. Economic models are an alternative analytical approach, directly targeted at forecasting the impact of interventions that were never implemented.

In this brief, we provide three examples on the use of economic modelling in the evaluation of development interventions. We use these examples to illustrate the application of modelling to explore various policy questions, and we then offer some tips on the use of modelling to evaluate development interventions. In doing so, we also show that models and experiments should be seen as complementary, rather than as alternative, approaches. Both have advantages and limitations, but their combined use enhances relevance for policy analysis. Over the last decade, many researchers have conducted randomised trials alongside economic models. The work of these researchers has shown that both approaches are strengthened by their combined use and the conclusions they lead to are full of policy implications.

This brief is based on the CEDIL Inception Paper No. 9, *Structural Modelling in Policy Making*, by Orazio Attanasio and Debbie Blair, and interested readers are encouraged to refer to this paper for exploring the arguments and examples used in this brief in more detail.

What is structural economic modelling?

Models are simplified representations of reality. In designing an economic model, we select a set of variables that we consider key determinants of a process. We then formally establish the relationships between these variables and we simulate what would happen if we were to change any of the variables considered. Modelling is common in the social sciences and not limited to economics. In addition to the already-mentioned epidemiological models, social scientists use *path analysis, directed acyclic graphs, agent-based modelling, system dynamics modelling, dynamic causal modelling, and several other approaches* (see Box 2 for a brief description of other modelling approaches). What characterises economic models, if you compare them with other models used in the social sciences, is their reference to a body of economic theory and a set of assumptions about the behaviour of economic agents. Economic models normally assume that people maximise their utility and that policymakers maximise social welfare.

Box 2: Other modelling approaches

Path analysis, and structural equation modelling more broadly, originated in the 1920s with the work of Sewall Wright. Path analysis and structural equation modelling consist of graphical representations of the causal relations between variables, which correspond to a system of mathematical equations. The diagrams can be very complex, representing numerous pathways between multiple observables and unobservable factors. Statistical methods are then applied to estimate the parameter values of the model. Structural equation modelling is often used in epidemiology to disentangle the effects of different impact pathways. For example, it has been used in the evaluation of water, sanitation, and hygiene interventions to uncover the relative impact of handwashing, water coverage, and sanitation in improving nutritional outcomes.

Directed Acyclic Graphs (DAGs) are associated with the work of Judea Pearl and originated in computer science. Like structural equation models, DAGs consist of graphical representations of the relationships between variables. In a DAG, the relations are causal and structurally stable, in the sense that the intervention will not cause the association to change. DAGs have gained some popularity in recent years and have helped researchers identify causal variables and relationships. DAGs are not a statistical method, but a model representation of causal relationships and can be applied to a large variety of problems. One limitation is their 'acyclic' character: they do not deal with simultaneous causes and therefore cannot model feedback loops, which are commonplace in the social sciences.

Agent-based modelling has a long history and has been used in various disciplines. It consists of computer simulations of interactions between agents that generate complex patterns. The reproduction of the complex patterns produced by flocking birds is a classic toy example of how the approach is able to model extremely complex phenomena starting from a few simple behavioural rules. The approach became popular in the social sciences by simulating interactions between individuals and their environment. Applications of agent-based modelling can be found today in the modelling of epidemics, financial transactions, migration, and waste management. Its application to the evaluation of interventions, however, has been very limited.

System dynamics, like agent-based modelling, was designed to understand the operation of complex systems. It was developed in the 1950s through the work of Jay Forrester and famously applied in the influential Limits to Growth report commissioned by the Club of Rome in 1972. In system dynamics, all elements of a system are interconnected and the connections are characterised by feedback loops. Concretely, system dynamic research employs two types of tool: causal loop diagrams and feedback models. The approach has not been widely employed in evaluation and it is best considered as a method that invites researchers to adopt a system perspective by delineating the nature of feedbacks loops and by identifying the ramified consequences of different courses of action.

Dynamic causal modelling is associated with the work of neuroscientist Karl Friston. Dynamic causal modelling is a framework to analyse the dynamic evolution of complex systems. The models are defined by systems of differential equations and are solved using Bayesian statistics. We are not aware of applications of this method to the evaluation of interventions. The approach came to prominence in 2020 thanks to some empirical applications to the modelling of the COVID-19 pandemic.

Maximisation of benefits can be defined to encompass benefits of others, and models developed by behavioural economists can accommodate non-rational behaviour. The attractiveness of economic models is that, by combining simple behavioural rules within relationships between variables, they are able to formulate hypotheses that would not have been possible by using simple intuition or observation.

As an example of a typical economic model, consider the classical model of school choice, which is popular in the economics of education. In this model, children—or more likely their parents—compare the costs of schooling (including books, school uniforms, fees, and missed work opportunities) with the benefits of schooling (the wage gain in accessing better occupations). Children will go to school if the perceived benefits over their lifetime outweigh the perceived costs. Now consider a cash transfer programme that remunerates parents for sending children to school. Viewed through the lens of an economic model, the programme is compensating for lost wage income from the children. If the comparison between benefits and costs is all that matters in making a schooling decision, the impact of the transfer on schooling will be similar to the impact of any other intervention that changes the balance of benefits and costs by the same amount. For example, the researchers may predict the impact of a reduction in school fees, or the effect of changing people's perceptions of the benefits of education.

Note also that understanding a cash transfer programme within an economic model of school choice highlights some potential impacts of the programme that are not immediately obvious. For example, the programme may not increase household incomes and consumption if the additional time in school would otherwise have been spent at work. Also, removing children from the job market may create pressures on local wages and discourage employment demand. In general, while some of these potential effects could be arrived at by intuition, others will not, and formal modelling has the power of making these potential effects clearly visible.

Economic models are often said to be 'structural'. What does this mean? They are structural in the sense that they are relatively independent of changes of circumstances and of the environment. For example, in the schooling model above, we did not specify whether the cash transfer was to be made to mothers or to fathers, whether it would be paid monthly or weekly, or whether it would be paid in cash or deposited in a bank account. A successful model should uncover the fundamental mechanisms behind people's behaviour. This behaviour should in turn be relatively independent of other factors. For example, how do people respond to price changes? How do parents make decisions about their children's education? The fundamental idea is that, if we are able to fully uncover these mechanisms, we can predict the outcomes in other contexts and the outcomes of different potential policy changes.

A structural model implies a good degree of simplification of the world. This is inevitable. You cannot give a full account of a complex social system. There are so many relevant details in a social system, and the best model of one ends up being the system itself. Any model is, therefore, a simplification. However, models can be tested with real data through a process of validation. If data are available from randomised experiments (or similarly credible quasi-experiments) on the impact of specific policies, the predicting power of models can be accurately tested. This is precisely the approach followed by all three examples illustrated below. The successful validation of a model through a comparison of its simulations and experimental observations increases the plausibility of the model and our confidence in its results. The use of models alongside randomised trials combines the best of both worlds.

There is a clear link here between economic modelling and the middle-range theory approach promoted by CEDIL. As economic models, middle-range theories are set to explain people's behaviour at a mid-level that lies between abstract generalisations (that are too general to be practical) and contextual concretisations (that are too specific to be useful). In this way, both approaches are able to make predictions of some general validity that transcend a specific context or a set of circumstances.

Policy applications of structural economic models: three examples

Free mosquito bed nets to fight malaria in Kenya

The advantages of economic models are better described through a series of examples. In our first example, Pascaline Dupas of Stanford University carried out a randomised trial in Kenya to assess the impact of a subsidy on the purchase of insecticide-treated bed nets to fight malaria. Individuals from six villages were given vouchers of different values that resulted in different bed net prices ranging from $0 to $3.80. The experiment showed that the subsidies increased demand substantially: 97% of those given the full subsidy purchased a bed net, but only 30% of those facing a price above $1.50 made a purchase. This study, along with others, ultimately convinced policymakers of the advantages of supplying mosquito bed nets free of charge. However, the argument at the time was against subsidising health products on the grounds that giving out free items would lead to waste and that offering a good at a discounted price would decrease its demand in the long term. This occurs because people accustomed to subsidised prices will no longer purchase the product at its market price once the subsidies are removed.

By running a randomised trial, Dupas showed that subsidies increased demand in the short term and that the free bed nets were not wasted. However, to investigate people's behaviour in the long term, she built an economic model. According to this model, people purchased nets if the perceived economic benefits were larger than the bed net price. People updated their perception of the bed nets' benefits through their use and by interacting with other bed net users.
The simulations suggested that the effect of the subsidy on the long-term willingness to pay is negligible. People's purchasing decisions at market prices are not influenced by being offered an item at a discounted price in the past. The simulation also showed that the effects of learning from experience and learning from others are large, thus pointing to a fundamental mechanism for the sustained use of bed nets.

The study concluded in favour of bed net subsidisation. In addition, the author used the model to infer the effect of other commonly subsidised health products such as cookstoves, water disinfectants, and deworming tablets. Drawing on the findings of the model, she inferred that a one-time subsidy on cookstoves would boost long-term adoption. On the other hand, a subsidy for water disinfectant would not have a significant impact, while a subsidy for deworming tablet would end up reducing demand in the long term.

Community-based monitoring and teacher incentives to increase schooling achievements in Rajasthan

In another example, Esther Duflo of the Massachusetts Institute of Technology, with colleagues Rema Hanna and Stephen Ryan, conducted a randomised trial in Rajasthan (in collaboration with the non-governmental organisation Seva Mandir) to address teacher absenteeism. Teacher absenteeism is common in the region and up to 24% of primary school teachers in India were found by a study to be absent during school hours. Seva Mandir had tried to reduce teacher absenteeism by offering monetary bonuses against attendance, by monitoring attendance through random visits, and through an ingenious system of photographs taken by the students themselves. The randomised evaluation found that the intervention had reduced absenteeism by more than 20 percentage points and, as a result, pupils' test scores improved significantly.

These results were important, but were also silent about the role of cash incentives and of monitoring in reducing absenteeism, while offering no guidance on the right size of teachers' incentives. To address these questions, the authors set up an economic model. According to the model, key factors such as income from teaching, leisure time, the probability of being fired as a result of missing days of work, and the threat of being publicly reprimanded affected whether teachers showed up to work and their overall labour supply. The results of the model suggested that the project operated predominantly through the financial incentive rather than by monitoring teachers. In addition, the authors were able to design an optimal teacher incentive scheme that would lower teacher absenteeism at the minimum cost. This optimal incentive turned out to be smaller than the one currently offered by Seva Mandir, thus allowing for immediate project savings.

Cash transfers to promote school attendance in Mexico

In a final example, Orazio Attanasio of the Institute for Fiscal Studies, with colleagues Costas Meghir and Ana Santiago, built an economic model to assess the impact of the PROGRESA cash transfer programme. PROGRESA is the most famous conditional cash transfer programme implemented in the late 1990s by the Mexican government in some of the poorest communities of the country. The programme gave mothers monetary transfers on the condition that their children attended clinic check-ups and nutrition education sessions, and more importantly on the condition that the children attended the final three years of primary school and the first three of secondary school. The impact of the programme has been analysed in dozens of papers and it has been found to be very successful at increasing school enrolment. Evaluations based on the experimental results of the programme, however, left many questions unanswered. Did the programme have an impact on non-targeted children? Would the programme be more effective if other children were targeted or if the transfers were differently allocated?

The researchers set out to address these questions using an economic model. According to the model, the decision to go to school was based on a comparison of the benefits and costs of continuing education. The model included variables such as expected returns to education, prevailing wages, the value of the grant, habits, and past schooling decisions. The authors found that the programme increased local wages by reducing local labour supply. This in turn created incentives for some children to leave school for work, thus partially tempering the impact the programme on school attendance. The authors then proceeded to simulate alternative policy interventions. These suggested that a budget neutral reallocation in favour of grants for pupils above Grade 6 would result in an increase in the attendance of secondary school pupils without affecting primary school attendance. The advice was contrary to government practice at the time of investing more resources in primary schooling. The researchers estimated that, by adopting their recommendations, the government would double the overall effectiveness of the intervention—a remarkable result for an intervention that was already successful. The model was also used to illustrate the effects of alternative interventions, for example a school building programme to reduce the maximum distance to secondary school to three kilometres. The authors found that the impact of such an intervention would be modest and not worth pursuing. As in the previous examples, these results could not have been derived from analysing the results of the randomised trial alone.

Key pointers on economic modelling

In this section, we discuss the strengths and limitations of structural models and provide pointers on their application.

Models can simulate the impact of different policies

Models can simulate the impacts of different courses of action. Policymakers are often confronted with situations in which several policy options are available to achieve a specific goal. These options cannot be tested one at a time until the right one is identified. Even if the policy choice is clear and unique, there are often infinite variations in which a specific policy can be implemented, and not all of them can be tested or empirically investigated. Economic models can predict the impact of modifying existing policies or of implementing entirely new policies. For example, a model that predicts how school attendance responds to monetary incentives can also predict how attendance will change if we were to change school fees, or if more schools were built and the walking distance to school were reduced. In the Mexican cash transfer programme example provided earlier, the researchers recommended reorienting the cash transfer programme in support of high school children. They confidently reached this conclusion without having to carry out another experiment and were similarly able to assess the impact of other potential interventions. This benefit of models is particularly useful when there is great uncertainty about the effectiveness of different courses of action, and testing different policies at the same time is clearly not an option.

Models can unpack complex interventions

Development interventions rarely consist of a single component with a single goal. They are often complex, consisting of multiple interacting components aimed at different goals at the same time. Development programmes are often implemented in the form of bundled interventions or packages. Impact evaluations can assess the effectiveness of the whole package but, depending on the design, they may be silent about the pathways of impact of the various activities. For example, an agricultural intervention may provide technical assistance to poor farmers as well as agricultural inputs such as fertiliser and seeds. Typically, evaluations assess the overall impact of the whole package but are unable to tell whether that impact is the result of technical assistance or of the inputs, or even of the two combined. Policymakers therefore receive no guidance with respect to what components of an intervention are more effective, or what combination of activities is most effective.

Randomised controlled trials can help in assessing bundled interventions by running multi-arm experiments where the different components are separately tested. This type of evaluation is feasible when the components are few, but it is prohibitively costly when there are many components. In these circumstances, economic models can help by teasing out the separate impacts of different intervention components without running multiple experiments. Economic models, by describing the pathways through which an intervention operates, are able to assess what specific project component, or what combination of components, is more important. Consequently, they are able to advise how the programme could be improved in this context. For example, in the school monitoring example in Rajasthan mentioned earlier, the researchers were able to conclude that the reduction in teacher absenteeism could be attributed more to the financial incentives than to monitoring from the bottom up.

Models can show how interventions work

Impact evaluations employing economic models aim at uncovering fundamental mechanisms behind social behaviour and can therefore uncover those 'middle-range' theories, which explain the effect of classes of interventions and across different contexts.
For example, the bed net study mentioned earlier showed that people do not become accustomed to subsidised prices as predicted by some, and that conversely their consumption choices are influenced by their experience in using a product and by learning from others. This insight allowed for predictions regarding the effectiveness of interventions that were not experimented. Since the purchase of different health products was affected in different ways by learning experience, the researchers were able to infer that a subsidy on a water disinfectant would not work, while a subsidised deworming tablet would be counterproductive. Note that this is not the same as claiming the results of economic models are externally valid in different contexts. The results of models are still obtained using data from a single population and from a specific context. They should therefore not be extrapolated mechanically to other contexts. In terms of external validity of the results, models do not appear to have a significant advantage over other evaluation methods.

Models encourage interdisciplinary research

Building an economic model requires deep knowledge and an understanding of the subject investigated. Studies relying on models are usually conducted by interdisciplinary teams. This has additional benefits of its own. Too often evaluations are conducted by researchers stuck in disciplinary silos, which can potentially limit the usefulness of the findings and their application in cross-cutting areas. As this pandemic has shown, multidisciplinary teams are able to address cross-cutting questions and are exploring new areas and solutions that are interlinked across thematic areas.

Meaningful models require engagement with stakeholders

As the statistician George Box said, we should never forget that 'all models are wrong, but some are useful'. All models are simplifications of a complex reality. Their results are always dependent on the assumptions made. In our approach, we propose to validate models by testing their predictive ability against the results of experiments, but the dependence on assumptions cannot be entirely eliminated.

A model is more likely to represent the context of operation of an intervention accurately, and more likely to simulate relevant policies, if multiple stakeholders are involved from the start and at different stages of the evaluation. Economic modelling is too often mainly a technical exercise, and successful modelling requires engagement with a different set of stakeholders. Early engagement with relevant stakeholders will contribute to the formulation of more meaningful models; it encourages a deeper understanding of the problem at hand, with all its ramifications, and it can make evidence more useful for informing decision making.

Models need investments in skills, time, and other resources

There are practical difficulties in building good models. Few researchers are trained in modelling and there is a steep entry cost in terms of acquiring the skills needed to computationally estimate a model. No software is readily available for setting up or solving a model. Building a good model requires time and resources.

Conclusion

This brief has outlined how economic modelling is designed to answer key policy questions that cannot always be effectively addressed by other methods. However, it does not suggest that economic modelling should replace experiments or other quasi-experimental studies. Theorising alone does not provide credible answers. On the contrary, **models should be used alongside existing impact evaluations.**

All our examples have referred to randomised controlled trials, but in principle data and results from quasi-experiments can also be used, provided they are of sufficient quality.

It has also been noted that building an economic model is not a simple exercise but requires considerable skills, time, and resources. This suggests it should not be carried out for each evaluation, but **only when we need to answer important policy questions**. Economic models should be promoted in those cases when our understanding of the impact of an intervention is limited; when we need to disaggregate the impact of different project components; and when we need to assess the effectiveness of different potential policies. This is preferably the case for large flagship interventions that need to be deeply analysed and understood.

Finally, structural models have the potential to foster interdisciplinary research and stakeholder engagement. A structural model is unlikely to provide useful insights unless these come as the result of a collaboration between researchers from different backgrounds. A technical team is unlikely to have deep knowledge of the context of an intervention and to know all the relevant policy options available. The engagement of multiple stakeholders should be sought to build meaningful models that generate useful evidence for decision making. Many fruitful collaborations in building interdisciplinary models, in partnership with policymakers, have recently emerged to provide support to decision makers in fighting the COVID-19 pandemic. Similar initiatives should be promoted in international development.

About this brief

This brief was prepared by Edoardo Masset with inputs from Radhika Menon, and is based on the CEDIL Inception Paper No. 9, *Structural Modelling in Policy Making*, by Orazio Attanasio and Debbie Blair.

The example applications described were taken from Dupas, P. (2014). Short-Run Subsidies and Long-Run Adoption of New Health Products: Evidence from a Field Experiment. *Econometrica*, 82(1), pp. 197–228; Duflo, E., Hanna, R. and Ryan, S.P. (2012). Incentives Work: Getting Teachers to Come to School. *American Economic Review*, 102(4), pp. 1241–1278; and Attanasio, O.P., Meghir, C. and Santiago, A. (2012). Education Choices in Mexico: Using a Structural Model and a Randomized Experiment to Evaluate PROGRESA. *The Review of Economic Studies*, 79(1), pp. 37–66.

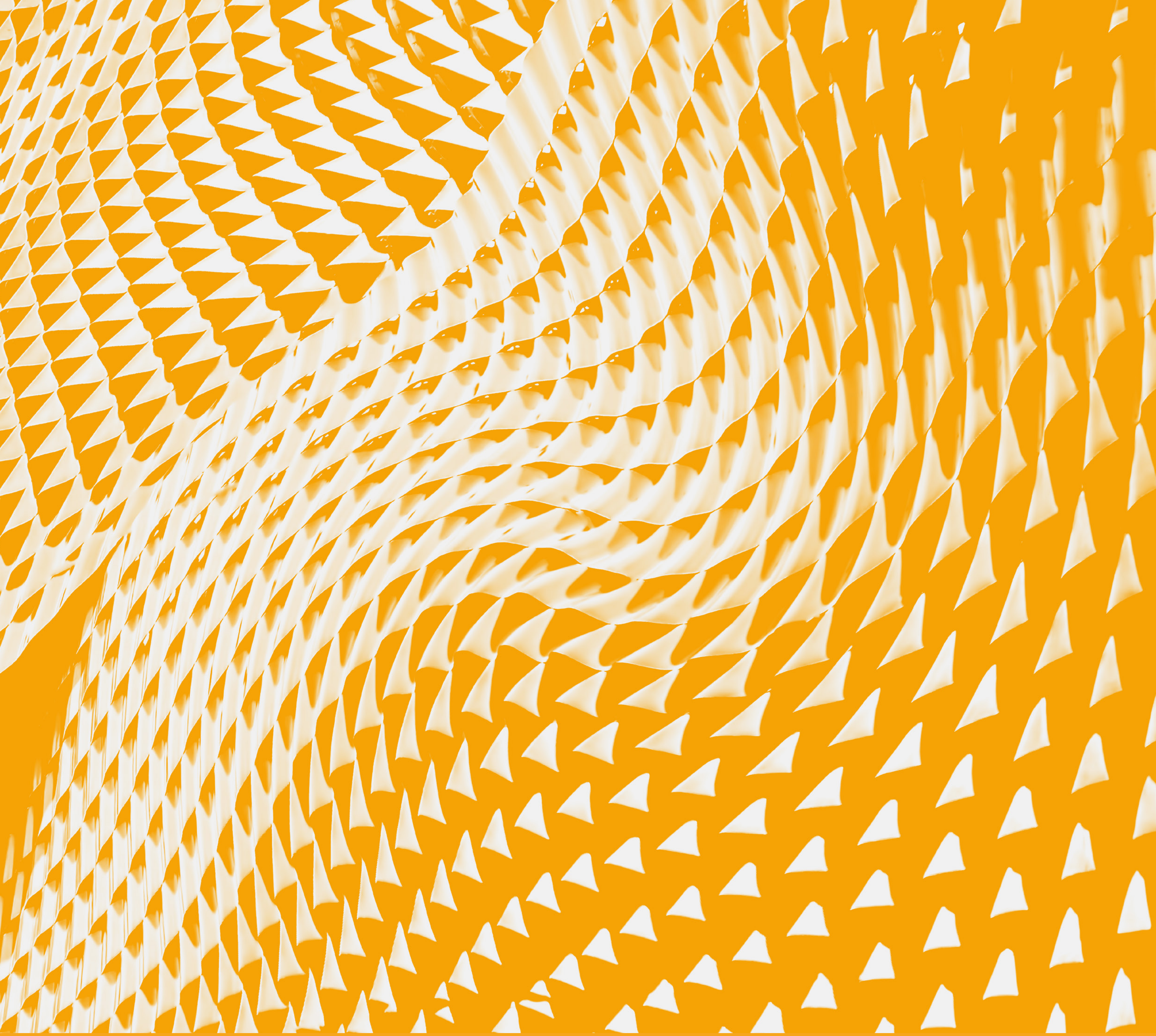
The following papers were also consulted and are recommended further readings: Low, H. and Meghir, C. (2017). The Use of Structural Models in Econometrics. *Journal of Economic Perspectives*, 31(2), pp. 33–58; Avery, C. et al. (2020). An Economist's Guide to Epidemiology Models of Infectious Disease. *Journal of Economic Perspectives*, 34(4), pp. 79–104; Galiani, S. and Pantano, J. (2021). Structural Models: Inception and Frontier. *SSRN Electronic Journal,* doi: 10.2139/ssrn.3799559; Carnap, T.V., Almås, I., Bold, T., Ghisolfi, S. and Sandefur, J. (2020). *The Macroeconomics of Pandemics in Developing Countries: An Application to Uganda*. CGD Working Paper 555. Washington DC: Center for Global Development; Heckman, J. (2010). Building Bridges Between Structural and Program Evaluation Approaches to Evaluate Policies. *Journal of Economic Literature,* 48(2), pp. 356–398; and Cartwright, N. (2020). Using Middle-Level Theory to Improve Programme and Evaluation Design. *CEDIL Methods Brief.* Oxford: CEDIL.

Suggested citation

Masset, E. 2021. 'Combining economic modelling and randomised controlled trials: An unexploited synergy', *CEDIL Methods Brief 3*. London and Oxford: Centre of Excellence for Development Impact and Learning. Available at : https://doi.org/10.51744/CMB3

Photo credits

p.3: Markus Spiske, p.6: U.S. President's Malaria Initiative, p.7: Michael Foley, p.8: World Bank, p.9: CGIAR Climate, p.11: World Bank.



Centre of Excellence for Development Impact and Learning

UKaid from the British people

About CEDIL

The Centre of Excellence for Development Impact and Learning (CEDIL) is an academic consortium supported by the UK Government through UKaid. The mission of the centre is to test innovative methodologies in evaluation and evidence synthesis and to promote evidence-informed development. CEDIL-supported projects fall into three programmes of work: evaluating complex interventions, enhancing evidence transferability, and increasing evidence use.

For more information on CEDIL, contact us at **cedil@opml.co.uk**
or visit our website **www.cedilprogramme.org**