AN ADAPTIVE TARGETED FIELD EXPERIMENT: JOB SEARCH ASSISTANCE FOR REFUGEES IN JORDAN

A. Stefano Caria
Department of Economics, University of Warwick

Grant Gordon
Essential

Maximilian Kasy
Department of Economics, University of Oxford

Simon Quinn
Department of Economics, University of Oxford

Soha Shami
Danish Refugee Council

Alexander Teytelboym
Department of Economics, University of Oxford

Abstract
We introduce an adaptive targeted treatment assignment methodology for field experiments. Our Tempered Thompson Algorithm balances the goals of maximizing the precision of treatment effect estimates and maximizing the welfare of experimental participants. A hierarchical Bayesian model allows us to adaptively target treatments. We implement our methodology in Jordan, testing policies to help Syrian refugees and local jobseekers to find work. The immediate employment impacts of a small cash grant, information and psychological support are small, but targeting raises employment by 1 percentage-point (20%). After four months, cash has a sizable effect on employment and earnings of Syrians. (JEL: C93, J6, O15)

Keywords: adaptive experimentation, refugees, active labour market policies, job search.

1. Introduction

Randomized controlled trials (RCTs) have become a widely-used method for policy evaluation (Duflo and Banerjee, 2017). In a conventional RCT, the designer randomly

assigns treatments to experimental subjects in order to precisely estimate the effects of all treatments. In many contexts, however, the experimenter is not merely interested in learning whether policies work. Instead, the experimenter wants to maximize the welfare of program participants. To do so, the experimenter only needs to learn which treatment works best. If the experimenter observes treatment outcomes over time, she can use this information in order to adaptively optimize treatment assignment for future experimental participants. The concern for participant welfare is becoming even more salient as the scale and scope of experiments grow, and when experiments are run with vulnerable populations (Banerjee et al., 2017a; Muralidharan and Niehaus, 2017; Haushofer and Metcalf, 2020).

Our first contribution is to introduce a methodology for adaptive targeted field experiments that balances the competing goals of precise treatment effect estimation (i.e., minimizing the variance of estimates of effects for *all* treatments) and maximizing the benefits to experimental participants. Our Bayesian algorithm has two key features. First, it is adaptive, i.e., it changes treatment assignment probabilities over time by incorporating information about the successes of treatments of existing experimental participants. Second, it is targeted, i.e., it uses information about the success rates of treatments in every group in order to target treatments for each individual group.

Our second contribution is to implement our methodology in a field experiment. As far as we know, ours is the first implementation of adaptive targeting in a field experiment in development economics. Our field experiment tested three active labour market policies for Syrian refugees and local workers in Jordan. We targeted treatments at 16 different strata of refugees and local workers. We find that our treatments had only a small impact on six-week employment outcomes of jobseekers, but did have larger impacts on longer-run outcomes. In particular, our results suggest that liquidity is an important barrier to labor market access for refugees. We also find that targeting increased employment by one percentage point (or 20 percent). We next describe our methodology, before turning to the context and details of our intervention.

Tempered Thompson Algorithm within a hierarchical Bayesian model. The first key feature of our methodology is that our treatment assignment is adaptive. The problem of adaptively assigning treatments in order to maximize outcomes during the experiment is known as a *multi-armed bandit* (MAB) problem; see Bubeck and Cesa-Bianchi (2012) for a review. MAB problems are often computationally intractable and a large literature in statistics has been devoted to finding tractable and effective heuristics to solve them. But MAB heuristics pose a problem for an experimenter interested in estimating the effects of all treatments: if the experimenter is quickly convinced that a particular treatment is *sub*optimal, she should stop assigning it in the future. As a result, the experimenter might miss out on learning about the effectiveness of good, though suboptimal, policies; furthermore, vanishing assignment shares for some treatments can also result in bias and size distortions for conventional inference (Hadad et al., 2019).

Our Tempered Thompson Algorithm combines the estimation objective of conventional RCTs with the welfare-maximizing objective of bandit algorithms. The

designer starts with a prior over the effectiveness of k different treatments; we recommend a diffuse and symmetric default prior. Every period, the designer observes the outcomes of some of the current participants in the experiment. As a result, the designer can estimate the posterior probability $\hat{p}_t^{dx}$ that treatment d is optimal for individuals from stratum x at time t. Then, at time t, the Tempered Thompson Algorithm assigns treatments in the following way, for individuals from stratum x:

With probability γ: **assign treatment d to individual i with probability $\dfrac{1}{k}$;**

With probability $(1 - \gamma)$:assign treatment d to individual i with probability $\hat{p}_t^{dx}$.

The Tempered Thompson Algorithm generalizes two classical treatment assignment protocols. When $\gamma = 1$, our algorithm boils down to a conventional randomized controlled trial. When $\gamma = 0$, our algorithm is the Thompson (1933) algorithm used in many online contexts, including platform revenue management, movie recommendations, and ad placement (Russo et al., 2018). However, when $0 < \gamma < 1$, the Tempered Thompson Algorithm (asymptotically) maximizes welfare of the participants subject to the constraint that every treatment has a probability of assignment at least $\dfrac{\gamma}{k}$. This allows the designer to target participant welfare while ensuring that she can learn something about the effectiveness of suboptimal treatments. One consequence of the lower bound in the assignment probabilities for each treatment is that we avoid the inferential problems for other adaptive algorithms as described by Hadad et al. (2019). Our main theoretical result (Theorem 1) formally establishes a trade-off between the welfare of participants and the precision of the estimates: as γ increases, the expected variance of treatment effect estimators falls, but the expected outcomes of participants also decrease. We also show that γ can be pinned down as solution to a decision problem of maximizing participant welfare subject to a constraint on frequentist inference (or, equivalently, on power or on standard error bounds). We discuss a number of additional considerations that may guide the choice of γ in the main body of the paper.

The second key feature of our methodology is that our adaptive assignment algorithm is targeted. We use a hierarchical Bayesian approach in order to estimate heterogeneous treatment effects. Our model allows us to learn the extent of effect heterogeneity across different strata. At each time period t, the treatment effect of each treatment d in each stratum x is estimated as a weighted average of the observed success rate for d in x and the observed success rates for d across all other strata. The weights are determined optimally by the observed amount of heterogeneity across all strata as well as the available sample size in a given stratum. The posterior probability $\hat{p}_t^{dx}$ that treatment d is optimal for stratum x is then calculated from this posterior distribution of treatment effects.

Implementation and Results. We implement our methodology in a field experiment designed to help Syrian refugees and local jobseekers in Jordan find wage work. Jordan is a highly relevant context in which to study employment policies for refugees, for at least two reasons. First, employment generation for refugees is a pressing policy concern. Since the start of the Syrian conflict, Jordan received

close to 700,000 Syrian refugees — one tenth of its original population (UNHCR, 2020a). Most Syrian refugees live in urban areas, outside camps, and 93% are below the national poverty line (Verme et al., 2015). Second, in 2016, Jordan and the International Community launched the Jordan Compact: in exchange for trade concessions and access to conditional financing, the Government of Jordan agreed to provide 200,000 work permits for refugees, lifting existing legal barriers that prevented them from obtaining work. The Jordan Compact has influenced refugee policy around the world (Betts and Collier, 2017). Legal restrictions to refugee employment are common in both developed and developing countries, and similar compacts are being launched to remove them, for example in Ethiopia. Jordan thus provides an opportune context to understand how to connect refugees to the new employment opportunities that are opening for them. To our knowledge, this is the first field experiment that studies the employment of refugees in a developing country context.

The experimental design and empirical analysis were specified before the start of the experiment in a pre-analysis plan submitted to the AEA registry.[1] The field experiment tests three types of support: a small, unconditional cash transfer (worth about one month of average monthly expenditure); information provision to increase the ability to signal skills to employers; and a behavioral nudge to strengthen job search motivation. These types of support correspond to three barriers – material, informational, and behavioral – that refugees and locals might face in finding and retaining jobs, and are often discussed in the qualitative and non-experimental literature (Schuettler and Caron, 2020). The program was implemented in Jordan by the International Rescue Committee (IRC) at the height of the Syrian refugee crisis.

In the experiment, we set $\gamma = 0.2$ in the Tempered Thompson Algorithm to ensure that in every period every one of three treatments and the control has at least 0.05 probability of being assigned. We define 16 strata: {Syrian, Jordanian} × {Female, Male} × {High school, No high school} × {Never employed, Ever employed}. Program intake started in mid-February 2019 and ended in December 2019. Overall, we sampled 3,770 individuals, approximately evenly split between Syrians and Jordanians. We track participants' employment outcomes with a 'rapid follow-up' phone interview six weeks after treatment; this survey asks simply about wage employment, and is used to implement our Tempered Thompson Algorithm . We then run detailed follow-up surveys two and four months after treatment. These surveys enable to us measures a broader set of impacts and to study effects over a longer time period.[2]

Our first finding is that, six weeks after being offered treatment, none of the interventions has significant or meaningful impact on the probability that individuals are in wage employment (the first outcome that we specified in our pre-analysis

---

[1] Available at https://doi.org/10.1257/rct.3870-2.2.

[2] We were unable to complete a six month follow-up interview due to the national lockdown in Jordan during the Covid pandemic.

plan). However, while the control-treatment difference in outcomes is close to zero, we estimate that the average impact of the optimised treatment (i.e. of offering the best possible intervention to each stratum) is about a 1 percentage point increase in employment (a 20 percent gain), suggesting some moderate short-term gains from targeting.

Second, we find that the cash intervention has large and significant impacts on refugee employment and earnings, two and four months after treatment. While employment rates remain stubbornly low in the control group, the cash grant raises job search rates and enables refugees to place more job applications. As a result, four months after treatment, the grant boosts employment by 3.8 percentage points (70 percent) and earnings by 65 percent. These are sizable impacts compared to those documented in the recent literature on active labor market policies (McKenzie, 2017). We also document substantial increases in hourly wages and in the probability of retaining a job between the two and four month interviews, indicating that match quality has also increased. Consistent with the existence of binding liquidity constraints, we find that these impacts are driven by individuals with below-median expenditure at baseline, and that baseline expenditure is significantly associated with job search intensity in the control group.

Third, the information and behavioral nudge interventions also boost job search among refugees and have significant impacts on employment and earnings after two months. However, these impacts are smaller than those of the cash grant and are ultimately short lived. Four months after treatment, we find weaker and insignificant impacts of these interventions.

Fourth, we find essentially no positive effects of treatment on the Jordanian sample.[3] The Syrians and Jordanians who take part in our study differ on multiple dimensions at baseline, and their outcomes evolve in different ways over time. In particular, the Jordanians have larger baseline expenditure, search at higher intensity and find jobs faster than the Syrians. This group may thus face weaker or different job search frictions, which are not addressed by our interventions.

These results shed light on the barriers to employment opportunities faced by refugees in a developing-country context. This evidence is particularly relevant for policy, as governments around the world consider expanding legal access to labor markets for refugees. In particular, our results point to the key role played by liquidity constraints, as in classical models of poverty traps (Banerjee and Newman, 1993; Balboni et al., 2020). Our comprehensive findings on these constraints – including the large employment impacts of a small unconditional cash grants, a strong control association between liquidity and job search intensity, and the large heterogeneity of treatment effects with respect to liquidity – gives clean evidence on the impacts

---

[3] We find a large and significant impact of job search from the psychological intervention, but no impacts on employment or earnings. Further, the cash and information interventions do not have significant impacts on any of our pre-specified outcomes.

of limited liquidity for refugees in a developing country.[4] At the same time, the job-search impacts of the other two interventions, which do not provide additional liquidity, show that cash is not a binding constraint for all refugees in our sample. Both information and motivation seem to further limit participation in labor markets.

Finally, we evaluate the performance of the Tempered Thompson Algorithm in our setting. We have three central findings. First, the Tempered Thompson Algorithm maximized six-week employment, an outcome that did not respond strongly to treatment. As a result, the proportions of individuals assigned to the various treatments did not depart very much from 25 percent and there were no large gains in average employment for our sample that came from adaptive randomization. Second, six-week employment responded to treatment more strongly for some specific subgroups and hence for these groups we see gains from adaptivity. For example, by the end of the trial we assigned 60 percent of the newly sampled Syrian women without tertiary education and work experience to the cash condition. Consistent with this, we find that the optimal targeted policy has a treatment effect on six-week employment that is one percentage point larger than the optimal non-targeted policy (though the confidence sets overlap). This provides suggestive evidence on the benefits of targeting. Third, we simulate what the performance of the Tempered Thompson Algorithm would have been had we targeted two-month employment. We find that the Tempered Thompson Algorithm quickly directs participants towards the optimal interventions and could have doubled the employment gains of a standard RCT. These simulations show that the Tempered Thompson Algorithm can generate substantial gains even with the levels of treatment effects we see in this context, but that the choice of the objective function is quite important. In the Appendix, we include a detailed discussion of several lessons for the design of adaptive trials that have emerged from our study, and some general guidance on extensions to the basic algorithm. The points we discuss include the use of surrogate outcomes, continuous (rather than binary) outcomes, choice of sample, strata and wave size, choice of prior, budget constraints, inference, non-stationarity, and alternative adaptive assignment algorithms.

Related literature. Our paper spans two distinct literatures. Methodologically, our work is related to experimentation, MAB problems, and targeted treatment assignment. While there is a large theoretical literature on optimal experimentation

---

[4] There is consistent evidence that interventions that provide much larger cash grants, worth up to one year of income, do not discourage work (Banerjee et al., 2017b). These interventions typically aim at fostering entrepreneurship, are often evaluated in rural contexts, and do not measure impacts on job search. Some related evidence comes from interventions that provide transport subsidies to facilitate the migration of the rural poor (Bryan et al., 2014; Meghir et al., 2019). Our findings are unique in that they study the effect of a much smaller grant, they document impacts on job search in an urban labor market, and they focus on a population of refugees. The only comparable intervention design is Banerjee and Sequeira (2020), who find a small unconditional cash grant boosts job search, but not employment, among a different population — young unemployed South Africans. Other urban studies such as Franklin (2018), Abebe et al. (2020) and Abebe et al. (2021) analyse the impacts of conditional transfers or financial application incentives that simultaneously relax individuals' budget constraint and decrease the 'price' of job search relative to other types of consumption. Thus, they do not offer conclusive evidence on the existence of binding liquidity constraints.

in MAB problems (e.g., Gittins (1979)), the bedrock of our analysis is the 'probability matching' algorithm due to Thompson (1933). Recently, a number of papers have shown for various settings that the Thompson algorithm is asymptotically efficient, among all possible adaptive algorithms, in terms of participant welfare (Agrawal and Goyal, 2012; Kaufmann et al., 2012; Agrawal and Goyal, 2013; Wager and Xu, 2021). While adaptive experiments are increasingly used in medical trials (Berry, 2006; FDA et al., 2018)[5] and commonly used in online advertising (Russo et al., 2018), we contribute to a growing number of papers in economics using these methods (Kasy and Sautmann, 2021; Kasy and Teytelboym, 2020, 2023). There is also a recent literature within economics on targeted treatment assignment both from non-Bayesian (e.g., Kasy (2015); Kitagawa and Tetenov (2018); Wager and Athey (2018); Chernozhukov et al. (2018)) and Bayesian perspectives (e.g., Dehejia (2005); Chamberlain (2011)). Finally, some recent papers consider the problem of assigning treatments under budget constraints both in the static setting (e.g., Bhattacharya and Dupas (2012)) and in the adaptive setting (e.g., Wang and Chen (2018); Kasy and Teytelboym (2023)). The papers that are closest to ours are Kasy and Sautmann (2021) and Kitagawa and Tetenov (2018). The Tempered Thompson Algorithm goes beyond the contribution of these papers in three ways. First, it addresses a different objective — that of balancing learning and welfare maximisation goals. Kasy and Sautmann (2021) focus on a learning objective, while Kitagawa and Tetenov (2018) focus on a welfare maximisation objective. In practice, experimenters have both learning and welfare maximisation goals, and our method enables them to choose how to balance these goals. Second, the Tempered Thompson Algorithm is both targeted and adaptive.[6] Third, previous demonstrations of adaptive experiments focused on online or on-the-phone interventions (e.g. Kasy and Sautmann 2021). Our implementation of Tempered Thompson Algorithm gives the first proof of concept that adaptive methods can be used for the delivery of complex policies in the field.

We also contribute to the literature on active labour market policies in developing and emerging economies. Specifically, ours is the first field experiment on employment of refugees in a development context.[7] The literature on active labour market policies has generally found that such policies have mixed effectiveness (McKenzie, 2017). This includes three novel experiments among educated youth in Jordan: one involving wage subsidy vouchers (Groh et al., 2016a), one involving training in soft skills (Groh et al., 2015b, 2016b), and one involving direct matching

---

[5] FDA et al. (2018) is a document by the US Food and Drug Administration which gives guidance to industry on how to implement adaptive trials. This is a very recent development, since the FDA has traditionally recommended simple, non-adaptive designs. An exception has been the use of early stopping rules, as part of ethical guidelines to protect participant welfare. If it becomes evident that a new treatment either severely harms all patients, or cures all patients, the trial has to be stopped. Such early stopping is the minimal form of adaptivity conceivable.

[6] Kasy and Sautmann (2021) discuss the possibility of targeting in the context of their algorithm, but do not implement this feature.

[7] Battisti et al. (2019) evaluate a job-matching intervention for recently-arrived refugees in Germany.

of job-seekers to firms (Groh et al., 2015a). However, in other contexts, recent experiments have identified several effective policy interventions: conditional cash transfers and financial application incentives have been found to increase short-term employment or applicant quality through increasing job search (Franklin, 2018; Abebe et al., 2020; Banerjee and Sequeira, 2020; Abebe et al., 2021), skill-signalling workshops can increase wages through improved assortative matching (Alfonsi et al., 2020; Bassi and Nansamba, 2020; Abebe et al., 2020), and detailed job-search plans have increased employment through more effective job search (Abel et al., 2019). We draw on each of these three recent areas of innovation to design our three treatments. The previous literature tends to focus on young nationals with poor attachment to the labour market (see, for example, Kluve et al. (2019)). Our work is novel in taking insights from those earlier experiments to a population of refugees, for whom constraints may be quite different.[8] In this way, our paper also relates to recent attempts to generalize experimental results across different contexts (see, for example, Meager (2019)).

Roadmap. The paper is organized as follows. Section 2 describes the humanitarian and the labour market context in Jordan, our sampling procedure, and the three treatments. Section 3 explains our adaptive treatment assignment algorithm and characterizes its theoretical properties. Section 4 presents our empirical findings. Section 5 assesses the performance of the Tempered Thompson Algorithm. Section 6 concludes. Appendix A formally establishes that the Tempered Thompson Algorithm is an approximate solution to the decision problem of maximizing participant welfare subject to a constraints on frequentist inference. Appendix B discusses lessons we learnt and recommendations for future adaptive field experiments including the use of surrogate outcomes, continuous (rather than binary) outcomes, choice of sample, strata and wave size, choice of prior, budget constraints, inference, non-stationarity, and alternative adaptive assignment algorithms. Appendix C gives the proof of the main theorem on the asymptotic properties of the Tempered Thompson Algorithm. Appendix D provides details on the Markov Chain Monte Carlo algorithm for the hierarchical Bayesian model, and Appendix E has some additional tables and figures. The Online Appendix contains treatment materials used in the field as well as additional tables and figures.

---

[8] Refugees often have lost substantial assets during displacement, are not familiar with the standards of their host labor market, and have suffered psychological shocks (Loiacono and Vargas, 2019; Schuettler and Caron, 2020). Additionally, language and legal restrictions to employment can be key barriers in some contexts (Battisti et al., 2019; Schuettler and Caron, 2020; Fasani et al., 2021), but are less salient in our setting: Syrian refugees in Jordan speak the local language and can work in blue collar occupations since the introduction of the Jordan Compact.

2. Context, sampling and treatments

The world is facing the largest refugee crisis since World War II, with over 80 million individuals displaced, about 26 million of whom are refugees (UNHCR, 2020b). Amidst this crisis, the duration of displacement has increased – with refugees now displaced for 10 years on average (Devictor and Do, 2017). The unprecedented magnitude and changing nature of displacement has catalyzed a big shift in thinking about how assistance is provided for refugees and internally displaced people.

Over the past decade, the international community has moved away from a model in which refugees are housed in camps – receiving aid in perpetuity – to a model focused on identifying sustainable solutions that integrate refugees and internally-displaced people into local communities and labor markets. In many contexts, this has fuelled a change from delivering basic commodities and food items to supporting individuals to gain access to employment.

Refugees face a unique set of constraints in accessing employment opportunities. They often lack the information, language skills and social networks needed to navigate labor markets effectively. Many have lost assets and have limited savings; this can constrain individuals from accessing the type of childcare, transit or basic needs required to get a job. Trauma, uncertainty and social exclusion may also reduce refugees' intrinsic motivation to search for an employment opportunity. These micro-level barriers may be compounded at the national level by governments who impose legal restrictions on whether or what types of jobs can be accessed.

2.1. The Syrian refugee crisis

Since 2012, the Syrian crisis has displaced more than 13.1 million people, making it one of the largest refugee crises of our time (UNHCR, 2020a). Approximately seven million are displaced internally within Syria; about another six million fled to neighbouring countries. The Government of Jordan estimates that, since the beginning of the Syrian crisis, nearly 1.3 million refugees have arrived in the country; of these, about 660,000 have registered with UNHCR (UNHCR, 2020a). Eight years into the conflict, Syrian refugees in Jordan face important needs for humanitarian assistance, for basic services, and for economic stability. In 2015, it was estimated that 93% of Syrian refugees in the country lived below the US$5 per day poverty line (Verme et al., 2015). At the same time, low-skilled Jordanians continue to suffer from pre-existing labor market challenges, including high unemployment, which leaves them also economically vulnerable (IRC, 2017; Government of Jordan, 2019; UNHCR, 2020a).

In an attempt to address this protracted displacement crisis, in 2016 the Government of Jordan and international donors signed the Jordan Compact. Under the Compact, donors pledged a total of US$2.1 billion in direct grants and US$1.9 billion in concessional loans to the Government of Jordan (Barbelet et al., 2018). The Compact also granted Jordan trade concessions that relaxed 'rules of origin' criteria and opened export markets in Europe. In exchange, the Government of Jordan

committed to important policy changes aimed at improving access of Syrian refugees to the labor market. In particular, the government agreed to ease administrative procedures to allow Syrian refugees to apply for work permits in manufacturing, agriculture, and construction – with a goal of providing work permits for up to 200,000 Syrian refugees (IRC, 2017).[9] The Jordan Compact has been hailed as an innovative approach for host countries and the international community to respond to protracted displacement crises.

2.2. The Jordanian labor market

The labour market in Jordan is characterised by very high levels of unemployment: a high proportion of the labor force is composed of individuals who do not have a job and are currently searching for one. In 2017, the unemployment rate among adult Jordanians was 18.5 percent – one of the highest rates in the Middle East and North Africa region. However, unemployment was even higher among Syrian refugees, reaching 25 percent (Tiltnes et al., 2019).

Search frictions are likely to decrease the returns to the job search efforts of this large number of refugee jobseekers.[10] First, refugees are often credit-constrained due to lost assets, networks, and sources of income (Government of Jordan, 2019). Second, they have little experience in and information on the formal labor market in the host economy, which could make it hard for refugees to convince recruiters of their employability. Third, they may experience self-control problems when it comes to searching for work, possibly resulting from the traumas of displacement (Shami, 2019). Importantly, these barriers are by no means unique to refugees in Jordan, as is consistently highlighted by the non-experimental literature on displacement (Loiacono and Vargas, 2019; Schuettler and Caron, 2020).

2.3. Sampling Syrian and Jordanian job-seekers

Our partner IRC enrolled individuals for a program called 'Project Match' on a rolling basis over a six-month period between February 10, 2019 and November 30, 2019. The program was active in three cities: the capital Amman, and the northern cities of Irbid, and Mafraq. To be eligible for this study, participants had to be: (i) Syrian refugees or Jordanian nationals with valid government identification, (ii) between 18 and 45 years old (inclusive), and (iii) willing to take up low-skilled formal wage work that pays approximately minimum wage in the immediate future. We verified that the

---

[9] In addition, the Government of Jordan also agreed to: (i) designate and develop five industrial zones, later called the Special Economic Zones (SEZs); (ii) allow Syrian refugees to formalize existing businesses and to set up new businesses; (iii) provide a small number of employment opportunities for Syrians in municipal works.

[10] Demand side constraints are also likely to play an important role. For example, firms often report difficulties in processing work permits for Syrians and refugees face strong competition from both Jordanian nationals and other migrants.

participants met these requirements and further collected information for the research during the intake registration interviews. At the end of the interview, participants were randomized into a treatment group based on the algorithm described in Section 3.

Participants were selected using a variety of passive and active recruitment methods. The passive methods entailed IRC employment service officers (ESOs) contacting potential program participants. We refer to this as 'passive' selection as it was initiated by the ESO and not by the program participant. In the majority of cases, employment officers learned about potential program participants from referrals given by community leaders, other programs or partner organizations, and other study participants. Additionally, the ESOs conducted door-to-door home visits to neighborhoods that were known to host a high number of refugees. These neighborhoods were identified using UNHCR maps and the experience of ESOs hired to work with Project Match. Further, individuals who had not been contacted by an ESO were also eligible to apply for the program. We refer to this as 'active' selection as it was initiated by the program participant. Individuals could enrol by visiting specific community-based organizations (CBOs), visiting the IRC offices, responding to ads posted on social media, or by attending an information session on Project Match at the UNHCR offices.

There were no major differences in the ways that Syrians and Jordanians were sampled. For both Syrians and Jordanians, the largest share of enrolments came from referrals, a passive sampling method. The second largest source of participants for both nationalities was enrolment by the job-seeker at a CBO (an active sampling method). Slightly more Syrians than Jordanians were sampled through home visits conducted by the ESOs. However, overall, low-skilled and more economically vulnerable Jordanians often resided in areas similar to those of refugees and also engaged actively with CBOs to access various forms of welfare. We summarise the frequency of these different sampling methods by nationality in Table A.2 in the Online Appendix. The proportion of participants enrolled through passive versus active methods changed over time, but not dramatically. In particular, in the months of May to July, 2019, more participants enrolled in Project Match through active methods. In subsequent months, this was largely reversed. We illustrate these patterns in Figure A.4 of the Online Appendix.

2.4. Key features of the sample

In total, we sampled 1,663 Syrians and 2,107 Jordanians. We report a number of descriptive statistics in Table 1. On several dimensions, the Syrian and Jordanian samples have similar characteristics. For both nationalities, 60 percent of the sample is composed of women, average age is about 29 years, and the average household is composed of about 5 individuals. Also, 2 percent of individuals of both nationalities are in wage employment and the average person has 5 years of work experience. Syrians, however, tend to be less educated on average (7 years vs 12 years). Our sample, while not intended to be representative as such, is nonetheless broadly similar to the near-representative JLMPS data in terms of respondents' age and education.

TABLE 1. **Descriptive statistics**

Sample	All	Syrian	Jordanian
	(1)	(2)	(3)
Female	0.60	0.60	0.60
Age	28.82	29.66	28.15
Household head	0.27	0.38	0.19
Household size	4.88	4.98	4.80
Education (years)	10.24	7.71	12.24
Spent at least 5 years in Jordan	-	0.95	-
Wage employed	0.02	0.02	0.02
Work experience (years)	4.48	4.99	4.10
Monthly HH Expenditure (JOD)	395.83	358.05	425.65
Searched for work	0.46	0.38	0.52
Applications (no.)	1.91	1.41	2.31
Hours search	3.14	2.60	3.58
Money spent (if searching)	8.03	6.90	8.69
Sample size	3770	1663	2107

Note: This table reports descriptive statistics for the Syrian and Jordanian samples at baseline. 'Monthly HH Expenditure' is total household expenditure in the last 30 days. This variable is obtained by summing reported expenditure in 9 different categories (including food, housing, utilities, water, transport, health, education, communication, and other). Missing values on specific expenditure items are replaced with the sample average. 'Searched for work' is a dummy for whether the person has done any job search in the last 30 days. 'Applications' is the number of job applications completed in the last 30 days. 'Hours search' is the number of hours of job search in the last seven days. 'Applications' and 'Hours search' take a value of zero if the respondent is not searching for work. 'Money spent' is the amount of money spent on job search in the last seven days, if the respondent is searching for work. 'Applications', 'Hours search' and 'Money spent' are winsorised at the 99th percentile of the nationality-specific distribution.

Many individuals in our sample live below the poverty line. Average household expenditure among Syrians is about 360 Jordanian Dinars (JOD) per month. This is approximately 72 JOD per person per month, or 2.4 JOD per day, which is well below the national poverty line of 3.5 JOD.[11] Daily expenditure among Jordanians is about 3 JOD – 25 percent higher, but still below the poverty line.

About half of the sample (40 percent of refugees and 50 percent of Jordanians) are actively looking for work at the time of the baseline interview. On average, refugees have placed about 1.4 job applications in the previous month and have spent 3 hours looking for work in the previous seven days. Among those refugees that were actively searching for work, total job search expenditure in the last seven days amounted to 6.9 JOD. Among Jordanians, average expenditure was 8.7 JOD.

We divide this sample in sixteen strata based on four dummy variables: (i) nationality (a dummy for whether the respondent is Jordanian, defined as having a Jordanian national ID); (ii) gender (a dummy for identifying as female), (iii) education (a dummy for having completed high school or more), and (iv) work experience (a dummy for having experience in wage employment). These strata form the basis of our

---

[11] This is a lower bound to daily per-capita expenditure, as in the calculation we assume that all members of the household are adult.

targeting strategy, discussed in the next section. In Figure A.5 of the Online Appendix, we show the distribution of observations across strata. While for most cells we have good sample sizes, we only have a small proportion of people, especially Syrians, who have education beyond high school.

As we discuss shortly, our treatment assignment algorithm targets assignment based on these covariates. We therefore do not expect these targeted covariates necessarily to be balanced across treatment arms. In Online Appendix Table A.3, we check for balance of a set of other key non-targeted covariates; that is, covariates not used for the stratification. To do this, we regress each covariate on the treatment dummies (with strata weights and robust standard errors). We find that we have good balance on all of those non-targeted covariates.

2.5. Treatments

On the basis of these key features, and working closely with local experts at the IRC in Amman, we designed three separate job search interventions.[12] Each intervention was designed to represent a distinct form of job search assistance, each having support in the recent empirical literature.[13] These interventions will be denoted by $D \in \{0, 1, 2, 3\}$ where 0 refers to respondents assigned to the control group; the three search interventions respectively provide cash, information, and psychological support. All interventions were delivered at the end of the intake interview or in the following seven days. In addition to these treatments, all respondents received 4 JOD (about US$5.60 USD at the time of the intervention), and an informational flyer with a very brief discussion of job interviews.

Control group. The control group received the 4 JOD and the basic informational flyer that were offered to everyone upon registration with Project Match.

Treatment 1: A labeled cash transfer. The cash support is a labeled cash transfer (LCT) of a value of 65 JOD (about US$92 at the time of the intervention). This transfer was intended to support the recipient to pay for the cost of job search – including transport, grooming, time costs and, for at least some study participants, childcare. It was designed based on evidence that small transfers cause large responses in job-search intensity (Herkenhoff et al., 2016; Franklin, 2018; Abebe et al., 2020). As shown in Table 1, the transfer enabled refugees to cover the average cost of job search for a period of about 10 weeks. The transfer was 'labeled' in that, at the time of distribution, study participants were offered recommendations on how they should use this cash, i.e., to help with the job search in the above-mentioned ways; however, respondents were also informed that they were free to spend the cash as they thought it was most appropriate (Benhassine et al., 2015). Upon delivery of the

---

[12] We prototyped and modified the interventions with about 130 respondents before commencing the randomized field experiment.

[13] Some additional respondents were also assigned to one of two separate 'direct placement' arms. These arms are not cross-randomized with the treatments that are analyzed in this study and are the focus of a separate paper.

intervention, participants received an empty ATM card, which was charged (within an average of seven working days) with a one-time cash payment of 65 JOD. Upon charging of the ATM card, recipients received an SMS notification. They also received an ATM guide pamphlet with a direct hotline number for reporting issues with cash withdrawal from ATMs. To put this intervention in context, it is important to note that this transfer is substantially smaller than the grants studied in the unconditional cash transfer literature. For example, Haushofer and Shapiro (2016) evaluate the impact of a transfer worth about two years of average consumption. In contrast, our transfer amounts to less than one month of average personal refugee expenditure.

Treatment 2: Information. The second intervention provided informational support. Prior evidence suggested that both Syrian refugees and low-skilled Jordanians had little understanding of either the interview process or the legal obligations owed by employers to their workers (Gordon, 2017). Specifically, respondents in this treatment received in-depth information on (i) how to prepare for and interview for a formal job in Jordan (following, in particular, the recent results from Abebe et al. (2020)), and (ii) the legal rights of employees in formal jobs. Information was delivered through face-to-face interaction with a trained Project Match employment service officer, two videos describing the formal jobs and associated labor laws from the eyes of a job-seeker, and two take-home paper tools.[14] The paper tools were designed for low-literacy participants and included cartoons for easy comprehension (see Online Appendix Figures A.1 and A.2). One of the tools was designed as an interactive myth-busting activity whereby participants are exposed to common myths about formal jobs and worker rights, and then upon scratching the surface of the box below the myth, can see the reality.

Treatment 3: Nudge. The third intervention is psychological support; we refer to this as the 'nudge' intervention. We provide a packaged intervention composed of (i) a four-week job-search planning calendar as in Abel et al. (2019) (see Online Appendix Figure A.3), (ii) an instructional video on how to use the calendar to plan for job search, (iii) a face-to-face demonstration delivered by the ESOs, and finally (iv) reminder SMSs. The instructional video begins with a statement on the potential impacts of job-search planning: 'Did you know that job-search planning can increase chances of finding work by up to 25%?' Additionally, if the participants send a text message to IRC with their job-search goals, they receive two reminder SMS, at the beginning of the week and at the end of the week. Through the calendar and the SMSs, participants track the number of jobs they intend to apply for and hours of job search they intend to spend, and can then check whether they met their goals for the week. This goal-setting and goal-reviewing exercise is designed to address self-control problems, which have been highlighted as a key constraint in recent research on job search (DellaVigna and Paserman, 2005; Caliendo et al., 2015; Abel et al., 2019).

---

[14] Participants did not receive any formal certificate from this or any of the other treatments.

2.6. *Follow-up surveys and attrition*

We measure the impacts of these interventions through three follow-up surveys, all administered by phone. First, we complete a very short follow-up survey six weeks after baseline. This survey is focused exclusively on measuring whether the respondent is currently in wage employment. The purpose of this survey was to provide the key outcome variable to implement the adaptive treatment assignment that we describe in the following section. For logistical reasons, this survey was conducted by a single enumerator, who had not previously interacted with respondents. We refer to this as the 'rapid follow-up survey'.

We then complete in-depth phone surveys two and four months after the baseline interview. (We had initially planned a six-month follow-up survey, but this was cancelled due to the Covid-19 pandemic and the strict lockdown measures imposed in Jordan.) We use these surveys to document the impacts of the program on a battery of outcomes specified in our pre-analysis plan. These interviews are carried out by the team of ESOs, with each respondent contacted by the same person who had interviewed the respondent at baseline and had enrolled them in the program. We refer to these two surveys as the 'ESO follow-up surveys'.

Our attrition is low throughout. In the rapid follow-up survey, attrition was 1.2 percent. In the ESO follow-up survey, we suffered attrition of approximately 2% at the two-month mark and 4.5% at the four-month mark (Online Appendix Table A.4). For Syrians, attrition in the nudge intervention is 1 percentage point above control at the two-month follow-up and 3 percentage points higher at the four-month point (Online Appendix Table A.5). For Jordanians, attrition in the information intervention is 1 percentage point above control at the two-month follow-up and 1.5 percentage points higher at the four-month point (Online Appendix Table A.6). Attrition for the cash intervention is very close to that of the control group (and not significantly different), in both follow-up surveys and both for Syrians and for Jordanians.[15]

3. Treatment assignment and inference

In this section we describe our treatment assignment algorithm. Our algorithm is a modification of Thompson sampling (Thompson, 1933; Russo et al., 2018). This modification is motivated by the fact that our experiment has two objectives. Our primary objective is to get as many experimental participants into formal employment as possible. In this context, employment is a reasonable proxy for welfare — and thus a sensible target — since all participants applied to participate in a program designed to help them find employment, so presumably finding employment is part of their

---

[15] We report *p*-values in Online Appendix Tables A.4, A.5 and A.6 by using randomisation inference, as described in Section 3.5. Some of the differences in attrition rates are marginally significant. Given the small rate of attrition overall, however, we consider it unlikely that our effect estimates discussed in Section 4 below are affected by differential attrition.

objective (see also Hussam et al. (2021) for recent experimental evidence on the social and psychological value of employment for refugees). Boosting refugee employment was also the goal of the organization that implemented the program.[16] Our secondary objective is to test the effectiveness of alternative interventions.

We implement our algorithm using a hierarchical Bayesian model; cf. Gelman et al. (2014). The data-generating process for the binary potential outcomes corresponding to each treatment d and stratum x is governed by a parameter θ^{dx}, which determines the success rate for this treatment and stratum. For a given treatment d, the parameters θ^{dx} are drawn from a prior distribution $Beta(\alpha^d, \beta^d)$. The hyper-parameters (α^d, β^d) determine the average success rate and dispersion of treatment effects across strata, for this treatment d. These hyper-parameters are assumed to in turn be drawn from a diffuse hyper-prior distribution π. The posterior distribution can be interpreted as follows. In every period, the experimenter observes treatment success rates for existing experimental participants for all strata and treatments. This allows her to learn the mean and dispersion of treatment effects across strata. She can then combine the estimate of these hyper-parameters with the observed success rate in a given stratum in order to calculate the posterior distribution of the success parameter θ^{dx} in that stratum. Finally, these posterior distributions can be used to calculate the probabilities $\hat{p}_t^{dx}$ that a given treatment is optimal for a given stratum. These probabilities are then used in the Tempered Thompson Algorithm.

After describing this Bayesian setup, we review Thompson sampling. Thompson sampling is based on the posterior probability that each of the treatments is optimal, conditional on observed covariates. We then introduce our modification, the Tempered Thompson Algorithm, which provides a compromise between Thompson sampling and full (balanced) randomization. In Theorem 1 we characterize how the Tempered Thompson Algorithm trades off our two objectives, helping participants and obtaining precise estimates. The source code for our assignment algorithm is available in a public repository.[17]

This section concludes with a discussion of inference. Our first method of inference is Bayesian. Our second method of inference uses p-values based on randomization inference. Randomization inference needs to take into account the adaptive and targeted form of treatment assignment in order to be valid.

We use the following notation. Let t denote the day the participant enrolled in the program and was offered the intervention and let i index individuals within days. Note that we have repeated cross-sections, not a panel, so that individual i on day t

---

[16] Our approach does not account for the benefits generated by the interventions in other domains – e.g. the quality of employment or the consumption gains deriving from the cash that is not spent on job search. However, in practice, it is unclear how these different dimensions should be weighted, since the experimenter does not generally observe the mapping between benefits and utility. To solve this problem, a related recent paper co-authored by one of us (Cesa-Bianchi et al., 2022) proposes an adaptive algorithm that enables experimenters to learn about participants' utility function and to simultaneously maximize their utility.

[17] At https://github.com/maxkasy/ThompsonHierarchicalApp. A corresponding interactive app is available at https://maxkasy.github.io/home/hierarchicalthompson/.

is different from individual i on day t' when $t \neq t'$. Let x index strata and d index treatments. Y_{it} is a binary outcome variable. Finally, m_t^{dx} denotes the total number of times that treatment d was assigned to individuals in stratum x up to time t, and r_t^{dx} denotes the corresponding total number of successes, that is, individuals for whom $Y_{it} = 1$.

3.1. Hierarchical Bayesian model

We consider a hierarchical Bayesian model with a data generating process, described by Eq. (1), and a prior, described by Eqs. (2) and (3) below. Let θ^{dx} be the average potential outcome for treatment d in stratum x. We assume that

$$Y_{it}^d | (X_{it} = x, \theta^{dx}, \alpha^d, \beta^d) \sim Ber(\theta^{dx}), \tag{1}$$

$$\theta^{dx} | (\alpha^d, \beta^d) \sim Beta(\alpha^d, \beta^d), \tag{2}$$

$$(\alpha^d, \beta^d) \sim \pi, \tag{3}$$

where (α^d, β^d) are the hyper-parameters and π is the hyper-prior (cf. Gelman et al. (2014, chapter 5)). Eq. (2) says that for a given treatment d, average potential outcomes θ^{dx} for all strata come from a Beta distribution governed by the hyper-parameters (α^d, β^d). Eq. (3) states that the hyper-parameters governing the distribution of average potential outcomes of each treatment across strata come from a common hyper-prior distributon π.

   We assume that the parameters $(\alpha^d, \beta^d, \boldsymbol{\theta}^{d\cdot})$ are independent across the treatment arms d. We choose a hyper-prior for the hyper-parameters (α^d, β^d) with a common density equal to $(\alpha + \beta)^{-2.5}$, up to a multiplicative constant. In doing so, we follow the recommendation of Gelman et al. (2014, p.110) for picking a "non-informative" hyper-prior.

   Intuitively, updating based on this prior works as follows. For each treatment d, we consider the success rates $q_t^{dx} = r_t^{dx}/m_t^{dx}$ across the different strata x. Based on these success rates, we learn the mean and dispersion of θ^{dx} across strata, as reflected in the hyper-parameters (α^d, β^d). Then we use these as a prior, which together with the cumulative successes r_t^{dx} observed for a given stratum x allows us to form an updated belief about θ^{dx} for that stratum.

   Denote by $\boldsymbol{\theta}, \boldsymbol{m}_t, \boldsymbol{r}_t$ the vectors of parameters, cumulative trials, and cumulative successes, where each of these is indexed by both d and x, and denote by $\boldsymbol{\alpha}, \boldsymbol{\beta}$ the vectors of hyper-parameters indexed by d. We sample from the posterior distribution of $(\boldsymbol{\theta}, \boldsymbol{\alpha}, \boldsymbol{\beta})$ given $\boldsymbol{m}_{t-1}, \boldsymbol{r}_{t-1}$ using the Markov Chain Monte Carlo algorithm described in Algorithm 1 in Appendix D.

3.2. Treatment assignment algorithm

Let p_t^{dx} denote the posterior probability that a treatment d is optimal in stratum x, in the sense that it maximizes the probability of employment. That is, define

$$p_t^{dx} = P\left(d = \arg\max_{d'} \theta^{d'x} | \boldsymbol{m}_t, \boldsymbol{r}_t\right). \qquad (4)$$

Equation (D.1) in the appendix shows how to estimate this probability by an average across Markov Chain Monte Carlo draws, which we denote $\hat{p}^{dx}$.

Two popular algorithms for assigning treatments in experiments are (i) fully random assignment, with equal probabilities across arms, and (ii) Thompson sampling. Our experiment is based on a combination of these two algorithms.

Fully randomized sampling assigns treatment d with probability $1/k$, where $k = 4$ is the number of different treatments, to units in every stratum. These assignment probabilities maximize power for tests of non-zero treatment effects. Thompson sampling, by contrast, assigns treatment d with probability $\hat{p}_t^{dx}$ to units in stratum x in time period t. Thompson sampling minimizes expected regret (cf. Agrawal and Goyal 2012; Bubeck and Cesa-Bianchi 2012), or equivalently maximizes average outcomes, in the large sample limit. As shown in these papers, it is in particular the case that expected regret only grows at a logarithmic rate with the number of experimental units. Russo and Van Roy (2016) prove worst-case bounds on the performance of Thompson sampling, using information-theoretic arguments.

Our primary goal is to maximize the labor market outcomes of experimental participants, but we also consider the precision of treatment effect estimates to be a secondary objective. Motivated by this combination of objectives, we assign treatment d to units in stratum x with probability

$$(1 - \gamma) \cdot \hat{p}^{dx} + \gamma/k, \qquad (5)$$

where γ is the share of observations that are randomized between treatment arms with equal probability. We refer to this procedure as the Tempered Thompson Algorithm.

In Appendix A, we show that γ can be pinned down as solution to a decision problem of maximizing participant welfare subject to a constraint on frequentist inference (or, equivalently, on power or on standard error bounds).

In our experiment, we measure employment outcomes Y_{it} only with a delay, six weeks after the intervention took place for each participant. As a consequence, treatment assignment is conditioned only on the outcomes of participants from six weeks before, or earlier. We assign participants in the first six weeks randomly to each treatment arm with probability 0.25.

3.3. Large sample properties

We now turn to a formal characterization of the large sample properties of our treatment assignment algorithm. We recapitulate and summarize our assumptions for

this characterization in Assumption 1. In the following, we use $\boldsymbol{\theta}_0$ to denote the fixed, true vector of average potential outcomes from which the data are generated. By contrast, we use $\boldsymbol{\theta}$ to denote the corresponding random vector which is drawn from the posterior distribution (belief) of the experimenter. The first step in Theorem 1 below, then, is based on the result that the posterior converges to the truth, that is, the distribution of $\boldsymbol{\theta}$ concentrates around $\boldsymbol{\theta}_0$.

ASSUMPTION 1 (Setup).   *Consider a fixed (non-random) $\boldsymbol{\theta}_0 = (\theta_0^{dx})$. Suppose that $d^{*x} = \arg\max_d \theta_0^{dx}$ is unique for all $x \in \{1, \ldots, n_x\}$, and denote $\Delta^{dx} = \theta_0^{d^{*x}x} - \theta_0^{dx}$. Assume that $(Y_{it}^1, \ldots, Y_{it}^k, X_{it})$ is i.i.d. across both i and t, and that*

$$Y_{it}^d | (X_{it} = x, \boldsymbol{\theta}_0) \sim Ber(\theta_0^{dx}).$$

Assume that $N_t \geq \underline{N}$ for all t and some constant $\underline{N}$, and that the prior distribution for $\boldsymbol{\theta}$ has full support.

Assume that treatment d is assigned to units in stratum x in period t with probability

$$(1 - \gamma) \cdot p_t^{dx} + \gamma/k,$$

where p_t^{dx} equals the posterior probability that treatment d is optimal in stratum x, and $0 < \gamma \leq 1$. Denote q_t^{dx} the cumulative share of observations assigned to treatment d in stratum x across the time periods $1, \ldots, t$, and p^x the probability that $X_{it} = x$.

The assumption that d^{*x} is unique holds generically.[18] We are now ready to state the main theorem.

THEOREM 1 (Large sample properties of Tempered Thompson Algorithm).   *Under Assumption 1, the following holds true as t (and thus $M_t = \sum_{t' \leq t} N_{t'}$) goes to ∞:*

1. *Consistency: The posterior probability p_t^{dx} that treatment d is optimal in stratum x converges to 1 in probability (conditional on $\boldsymbol{\theta}_0$) for $d = d^{*x}$, and to 0 for all other d.[19]*
2. *Converging shares: The cumulative share q_t^{dx} allocated to treatment d in stratum x converges in probability to $\bar{q}^{dx} = (1 - \gamma) + \gamma/k$ for $d = d^{*x}$, and to $\bar{q}^{dx} = \gamma/k$ for all other d.*
3. *Converging regret: Average in-sample regret,*

$$Regret_t = \frac{1}{M_t} \sum_{i,t} \Delta^{D_{it} X_{it}}$$

---

[18] The case where the uniqueness assumption is close to being violated is discussed in Section 3.4 below.

[19] Note that this statement refers to frequentist consistency (given $\boldsymbol{\theta}_0$) of a Bayesian posterior probability (which averages over $\boldsymbol{\theta}$).

converges in probability to

$$\gamma \cdot \frac{1}{k} \sum_{x,d} \Delta^{dx} \cdot p^x.$$

4. *Converging estimator: The normalized average outcome for treatment d in stratum x,*

$$\sqrt{M_t} \left(\bar{Y}_t^{dx} - \theta_0^{dx} \right),$$

converges in distribution to

$$N \left(0, \frac{\theta_0^{dx}(1 - \theta_0^{dx})}{\bar{q}^{dx} \cdot p^x} \right).$$

Furthermore, the normalized average outcomes for different treatments and strata converge jointly to a multivariate normal with covariances equal to 0.

The large sample result of Theorem 1 characterizes the trade-offs in choosing γ. The parameter γ allows us to interpolate between non-adaptive, conventional randomization ($\gamma = 1$) and Thompson sampling ($\gamma = 0$). The former is optimal for minimizing the expected variance of all treatment effect estimators. The latter is optimal for minimizing the expected regret (maximizing expected welfare) for the participants in the experiment.

As we increase γ, starting from a value of 0, the expected in-sample regret increases linearly in proportion to γ. On the other hand, the asymptotic variance of conditional average treatment effect estimators, comparing the conditionally optimal treatment to its alternatives, is given by one over the total sample size, times

$$\frac{\theta_0^{d^{*x}x}(1 - \theta_0^{d^{*x}x})}{((1 - \gamma) + \gamma/k) \cdot p^x} + \frac{\theta_0^{dx}(1 - \theta_0^{dx})}{(\gamma/k) \cdot p^x}.$$

This number is decreasing in γ, since higher γ means a more balanced distribution of observations across treatment arms. In our application, we trade off these conflicting objectives by setting the share of observations for which treatment is fully randomized to $\gamma = 0.2$, which implies that the probability of being assigned to each treatment is bounded below by 0.05. Given that the baseline employment rate is 5% (see Table 2 below), our choice of γ then means that in the large sample limit (i.e., when only 5% of the sample in assigned to some treatment) the standard error of a treatment arm estimate would be bounded by

$$\sqrt{\frac{E[Y|D=d]}{N \cdot (\gamma/k)}} \approx \sqrt{\frac{0.05 \times (1 - 0.05)}{3668 \cdot (0.2/4)}} = 0.016,$$

which is about twice the standard error we would get under a fully randomized design. Given that in practice we do not see a full convergence toward this large sample limit, our standard errors are actually smaller.

More generally, the relative weight that researchers should assign to participant welfare versus estimator precision might depend on a number of contextual factors. One factor is the size of the experimental population relative to the population that will be ultimately reached by interventions informed by the trial. The larger this size, the greater the weight one might wish to put on estimator precision, and the larger the optimal γ. Another factor is whether the treated population is particularly vulnerable and hence requires special protection. If that is the case, ethical considerations suggest a greater weight for participant welfare, and thus a smaller optimal γ.

3.4. Discussion of Theorem 1 and the Tempered Thompson Algorithm

Several observations are worth making about the properties of the Tempered Thompson Algorithm and Theorem 1. First, the theorem implies that the large sample properties of the Tempered Thompson Algorithm do not depend on the prior (as long as the latter has full support): In large samples the data dominate the prior, the posterior is consistent, and thus assignment shares become independent of the prior. Relative to pure Thompson sampling, this happens even faster for the Tempered Thompson Algorithm with $\gamma > 0$.

The flip side of this large-sample robustness to the prior is robustness to the data in the initial periods, for three distinct reasons. First, Bayesian inference optimally combines data and prior, and therefore down-weights outliers among the initial observations. This stabilizes assignment shares in initial periods, and makes them closer to an equal division among treatment arms. Only when evidence has accumulated that some treatment arms are better than others do assignment shares become unequal. Second, relative to Thompson sampling the Tempered Thompson Algorithm additionally shrinks assignment shares toward the balanced assignment. And third, the outcomes in our setting are bounded, and therefore the influence of any single observation on the posterior is necessarily bounded, as well.

The properties of the Tempered Thompson Algorithm listed in Theorem 1 rely on having sufficient power (enough observations) to be able to distinguish the best treatment with high probability. It therefore presupposes that learning is sufficiently fast. An alternative characterization might consider the worst-case behavior of the Tempered Thompson Algorithm across possible values for the parameter vector $\boldsymbol{\theta}_0$, for a given T. As it turns out, the worst-case behavior in terms of in-sample regret is driven by intermediate parameter values which are such that the treatment effects Δ^{dx} (relative to the optimal treatment) are of the same order of magnitude as the standard errors for estimates of these treatment effects, that is, of order $1/\sqrt{T}$ (Bubeck and Cesa-Bianchi, 2012; Russo and Van Roy, 2016). If treatment effects are larger, the best treatment is discovered quickly and in-sample regret is low. If treatment effects are smaller, it doesn't matter as much which treatment participants are assigned. The behavior of algorithms such as the Tempered Thompson Algorithm can also be characterized using local-to-zero asymptotics, as in the recent pioneering work by Wager and Xu (2021). A formal characterization of the worst-case and local-to-zero behavior of the Tempered Thompson Algorithm is left for future work.

3.5. Inference

One concern about adaptive experimental designs is that they lead to biased inference (see, e.g., Hadad et al. 2019). Part 4 of Theorem 1 implies, however, that this is not the case for the Tempered Thompson Algorithm in large samples. Sample averages in each treatment arm are consistent, asymptotically unbiased, and normally distributed, so that inference can proceed as if treatment assignment were not adaptive. This is true because assignment probabilities for each arm in each stratum are bounded away from 0 when $\gamma > 0$. In our empirical analysis, we nevertheless consider two methods for inference that do not rely on such asymptotics, but instead are exactly valid in finite samples despite adaptive assignment, as detailed next.

Our first form of inference is Bayesian, based on the hierarchical default prior described in Section 3.1 above. To construct credible sets (i.e., sets that have a given posterior probability of containing the true parameters), we report 0.025 and 0.975 quantiles, based on Markov Chain Monte Carlo draws. We do so for all our estimates listed in the previous section. This yields sets that have a posterior probability of 95% to contain the true parameters, conditional on the data of the experiment.

We would like to emphasize that standard Bayesian inference remains valid in finite samples for adaptive designs such as ours, since the likelihood function is not affected by adaptivity. In large samples, as long as $\gamma > 0$, and as long as the optimal treatment is unique in each arm, we conjecture that our credible sets also have 95% frequentist coverage probability, i.e., they are confidence sets in the usual sense; cf. van der Vaart (2000), chapter 10. This relies on the fact that the share of observations assigned to each treatment in each stratum is bounded below, asymptotically.[20]

Additionally, we provide randomization-based p-values that are valid under the sharp null hypothesis that there are no treatment effects, i.e., under the null that $Y_{it}^d = Y_{it}^{d'}$ for all d, d', i, t. Under this null, we can generate counterfactual data by re-running our assignment algorithm repeatedly, leaving outcomes as they are in our data, but generating new treatment assignments. The distribution of test-statistics over this re-randomization distribution can be used to construct critical values and p-values that are exact in finite samples, under the sharp null.

4. Results

In this section, we discuss the impact of the interventions and the performance of the Tempered Thompson Algorithm. We first present a set of results on wage employment based on the rapid follow-up survey carried out six weeks after baseline. We do this in two different ways. First, we present Bayesian posteriors and credible sets. Second, we report the difference between weighted average employment in each treatment

---

[20] A formal proof is left for future work, but should follow from item 4 of Theorem 1.

group and in the control group.[21] Here, we use randomization inference to construct a p-value of the sharp null of no treatment effect.[22]

We then present a broader set of results from the longer surveys we carried out two months and four months after baseline. These surveys capture wage employment, but also measure job-search, earnings, well-being, social integration and migration intentions. For each outcome, we report weighted averages and randomization inference p-values as explained above.

We present an evaluation of the performance of the Tempered Thompson Algorithm in our setting, including three 'welfare contrasts' that quantify the overall impact of our interventions. We also simulate the counterfactual performance of the Tempered Thompson Algorithm with two-month employment as its objective.

Finally, we summarize the quantitative and qualitative evidence on the main barriers to job search among refugees.

4.1. Employment in the rapid follow-up survey

Job-finding rates in the control group are consistently low, especially for Syrians. Six weeks after joining the program, the average control wage-employment rate is 4.9 percent (Table 2). Further, individuals sampled at different points in time tend to have similar six-week employment rates, except for somewhat higher rates for those sampled in the first month of the experiment (Figure A.7 in the Appendix). These averages, however, mask substantial heterogeneity (Table A.10 in the Online Appendix). Employment rates among Jordanians (6.8 percent) are more than twice as large as employment rates among Syrians (2.7 percent). Similarly, the male employment rate (7.7 percent) is more than twice as large as the female employment rate (3.1 percent). Overall, most subgroups have employment rates below 10 percent.[23] Given that job search at baseline was substantial, this highlights the difficulty of finding work in this labour market.

---

[21] Weighting is necessary as the samples in each experimental group are mechanically unbalanced due to our adaptive randomization procedure. We report *weighted averages* of the form:

$$\beta_j^d = \frac{1}{N} \sum_{it} \frac{\mathbf{1}(D_{it} = d)}{a^{dx}} \cdot W_{it}^j,$$

where

$$a^{dx} = \frac{\sum_{it} \mathbf{1}(D_{it} = d, X_{it} = x)}{\sum_{it} \mathbf{1}(X_{it} = x)}.$$

W_{it} is some outcome of interest of individual i sampled on day t, D_{it} is the treatment status of this individual, X_{it} is the stratum, and N is the total number of experimental participants. So a^{dx} is the overall proportion of individuals in stratum x assigned to treatment d by the end of the experiment.

[22] We only compute randomization inference p-values for individuals that were contacted during the rapid follow-up survey. For consistency, we thus drop observations that have attrited in the rapid follow-up survey, but not in the two and four month follow-up surveys (39 and 24 observations, respectively).

[23] In Table A.9 in the Online Appendix, we look at the full break-down in sixteen strata, we find that three strata have employment rates above 10 percent. However, in two of these case, the strata have very few observations and so our measure of employment rate is likely to be noisy.

TABLE 2. **Impacts on employment in the rapid follow-up survey**

	All	*Syrians*	*Jordanians*
Cash	0.006 (0.296)	0.013 (0.123)	-0.001 (0.531)
Information	-0.005 (0.690)	-0.004 (0.626)	-0.006 (0.648)
Nudge	0.003 (0.388)	0.005 (0.348)	0.002 (0.463)
Control mean	0.049	0.027	0.068
Observations	3668	1633	2035

Note: The table reports impacts on wage employment at the time of the rapid follow-up interview (6 weeks after treatment). Next to each treatment effect estimate, we report a randomization inference p-value, obtained using the procedure discussed in Section 3.5. The first panel reports results for the whole sample. The second panel reports results for Syrians. The third panel reports results for Jordanians.

Our main finding is that, six weeks after the start of the program, none of the interventions increase employment for the average program participant. We report Bayesian posteriors on the impacts of the different treatments and the respective credible sets in Figure A.8 in the Appendix. These posteriors indicate that the impact on employment is always smaller than 1 percentage point. We confirm this result by reporting differences in weighted employment rates in Table 2.

We are unable to find evidence of treatment impacts for specific, pre-specified groups of individuals. In Figure A.8, for example, we show treatment effects after splitting the sample by nationality and do not find evidence of impacts on employment on either Syrians or Jordanians. Posteriors are somewhat larger for Syrians than for Jordanians, but the credible sets always overlap. We report credible sets for all sixteen strata in Table E.A.1 in the Appendix. Further, Table 2 above and A.10 in the Online Appendix report differences in weighted employment by group which confirm these findings. Employment effects are somewhat larger for Syrians (e.g. employment rates in the cash group are a 1.3 percentage point higher than in the control group), but these effects are not significantly different from zero.

4.2. Estimated impacts in the ESO follow-up surveys

We now turn to explore effects at the two-month and four-month follow-up points, using detailed data from the follow-up surveys conducted by the Employment Service Officers (ESOs).

4.2.1. Impacts on job search. Despite the null impacts on employment after six weeks, we document that all interventions generate marked increases in job search among Syrians.[24] As shown in the first panel of Table 3, the cash transfer raises the

---

[24] The analysis of job search outcomes in this subsection was not pre-specified. In the Pre-Analysis Plan we committed to studying the impacts of the interventions on five main longer-term outcomes (which we report in the next subsection). Further, we anticipated that, motivated by our main results on those outcomes, we would run a number of additional exploratory regressions to better understand treatment mechanisms. This subsection presents this exploratory analysis.

TABLE 3. **Job search impacts after 2 months**

	Searched for work	Applications	Hours job search
Syrians			
Cash	0.056 (0.077)	0.518 (0.043)	0.794 (0.133)
Information	0.047 (0.123)	0.423 (0.072)	0.056 (0.482)
Nudge	0.037 (0.195)	0.648 (0.016)	0.698 (0.157)
Control mean	0.437	1.19	4.162
Observations	1536	1440	1444
Jordanians			
Cash	0.033 (0.165)	-0.055 (0.553)	-0.672 (0.831)
Information	0.025 (0.255)	-0.501 (0.874)	-0.457 (0.757)
Nudge	0.065 (0.030)	0.458 (0.130)	0.241 (0.350)
Control mean	0.577	2.71	5.792
Observations	1897	1783	1717

Note: This table reports treatment effects on the three variables measuring job search, 2 months after the baseline interview. 'Job search' is a dummy for whether the person has done any job search in the last 30 days. 'Applications' is the number of job applications completed in the last 30 days. 'Hours search' is the number of hours of job search in the last seven days. 'Applications' and 'Hours search' are windsorised at the 99th percentile of the nationality-specific distribution. The first panel reports impacts for Syrians, and the second panel reports impacts for Jordanians. Next to each treatment effect, we report a randomization inference p-value.

proportion of Syrians who look for work two months after baseline by 5.6 percentage points (a 13 percent increase over a control job-search rate of 44 percent) and leads Syrians to place 0.5 more job applications (a 40 percent increase over a control mean of 1.2 applications). Similarly, the information intervention and the nudge intervention raise job search rates by 4.7 percentage points and 3.7 percentage points respectively. Both of these interventions also have significant impacts on job applications: a 35 percent increase for the information intervention and a 55 percent increase for the nudge intervention.

Among Jordanians, the cash and information interventions have smaller and insignificant impacts on job search (Table 3, second panel). For example, the cash intervention is associated with a 3.3 percentage point insignificant increase in the job-search rate. However, we find a large and positive impact of the nudge intervention on job search (6.5 percentage points). Importantly, no intervention is associated with a significant impact on job applications among Jordanians and, for both the cash and information intervention, the effect is actually negative. Finally, the second panel of Table 3 also highlights that Jordanians search much more intensely than Syrian refugees in the absence of the interventions: the control job search rate is 30 percent higher and the control number of job applications is twice as large.

4.2.2. Impacts on labour market outcomes. We find that both the cash and information intervention improve Syrian refugees' labour market outcomes at the

two-month and four-month marks.[25] We report the relevant coefficient estimates and randomisation inference p-values in Table 4 and Table 5. Offering cash leads to a significant increase in the employment rate of more than 50 percent (an effect of 5.2 percentage points in the two-month survey and 3.8 percentage points at the four-month survey) and a significant boost in earnings of about 40 percent after two months and of 65 percent after four months.[26]

We estimate that the information intervention increased employment and earnings by almost the same amount as the cash grant two months after baseline; four months after baseline, this intervention generates a 40 percent increase in employment (i.e., 1.9 percentage points, marginally insignificant) and a 55 percent increase in earnings (again, marginally insignificant). We also find that the nudge intervention has weaker and short-lived effects on the labour market outcomes of refugees. Four months after baseline, we are unable to document any significant impacts of this intervention.

Note that we find different results – in both control group employment rate and in treatment effects – when we compare the six-week rapid follow-up surveys with the two-month ESO follow-up surveys. We attribute this discrepancy to the fact that the six-week survey was conducted by a different enumerator (not known to the respondent), in a much shorter questionnaire. As discussed in Section 2.6, ESO surveys were conducted by staff members familiar to the respondents, whereas the rapid-response surveys were conducted by someone whom the respondents had not previously met. It is thus likely that many respondents, and particularly Syrian refugees, had more trust in the ESO enumerators, and hence were more likely to discuss informal work with them. (Due both to ethical and confidentiality concerns, we did not ask respondents – either in the six-week survey or subsequently – to specify whether a particular wage job was formal or not.) In support of this interpretation, we show in Table A.15 in the Online Appendix that a substantial share of the increase in wage employment documented in the two-month follow-up survey occurs in jobs that pay just below the formal monthly minimum wage: a useful proxy of informality. In Section 4.3, we show that these jobs have high hourly earnings, but fewer hours worked per month than formal jobs (and so, overall, lower monthly pay). During our qualitative interviews (Section 4.3.3), many respondents expressed a preference for this type of work.

The magnitude of the effects on refugee employment and earnings that we document is large relative to the estimates reported in the recent literature for other active labor market policies in developing countries (McKenzie, 2017). In proportional

---

[25] The analysis in this subsection was pre-specified. We summarise all variable definitions in Table A.12. We report results disaggregated by nationality in Table 4 and 5. We report results for the full sample in Table A.13 and Table A.14 in the Online Appendix.

[26] We compute a set of sharpened q-values to control for the fact that we test for impacts on each outcome three times – once per intervention. The q-values for the impacts of the cash intervention on employment after 2 and 4 months are, respectively, 0.051 and 0.081. The q-values for the impacts of the cash intervention on earnings after 2 and 4 months are, respectively, 0.12 and 0.138.

terms, both the earning and the employment effects are at the top of the distribution of the estimates reported in the literature. While this is partly driven by the low control employment rate, in absolute terms, the employment effect is still close to the top of the distribution of existing estimates for job search assistance policies (but smaller than the impacts of the most effective wage subsidy and training interventions). If the four-month impacts were sustained for about one year, the increase in earnings would equal the size of the cash grant. In comparison to the existing literature, however, our results speak only to a relatively short impact duration (namely, two-months and four-months after treatment); as explained earlier, the Covid lockdown forced the cancellation of our planned six-month follow-up survey (and, of course, caused massive disruption to the Jordanian labour market more generally).

For refugees, the cash grant and information interventions are also associated with small, insignificant increases in the well-being index and with an insignificant 7 percent drop in the proportion of people that intend to migrate outside of Jordan. While the migration effect is not significant, in absolute terms, the 4 percentage points decrease we estimate is commensurate to the size of the positive employment effects of these interventions. Finally, we do not find any impacts on social integration, consistent with recent evidence suggesting that baseline social integration for Syrian refugees in Jordan is high relative to the experience of Syrian refugees in European countries or in the US (Alrababa'h et al., 2019).

For Jordanians, on the other hand, we are unable to find evidence of labour market impacts for any intervention (Table 4 and Table 5). This is particularly surprising for the nudge intervention, which has positive impacts on job search for this population. Further, for the cash intervention, we document an (insignificant) 2.9 percentage point reduction in employment and an (insignificant) 25 percent reduction earnings after four months, but a contemporaneous, significant increase in the well-being index of 0.06 of a standard deviation. This may be consistent with the cash intervention enabling jobseekers to reject offers for undesirable jobs.

Finally, we test for potential impacts on the margin of self-employment for Syrians.[27] We find – at best – small impacts of treatment on the probability of self-employment; we report this regression in Online Appendix Table A.16. Similarly, in the same table, we show that treatment effects on total personal income are very close to our estimated treatment effects on wage income, and statistically indistinguishable from those earlier estimates; this is consistent with a small treatment effects on self-employment income. These results are consistent with two pieces of contextual evidence. First, for the Syrian refugee sample, there are legal restrictions on self-employment (UNHCR, 2017). Specifically, they can only operate in a limited number of sectors (food processing, handicrafts/tailoring, tourism, manufacturing, and agriculture). Second, the value of our transfer was unlikely to be sufficient to start a

---

[27] This analysis was not pre-registered; we thank an anonymous referee for encouraging us to discuss this.

business in this context.[28] In sum, the self-employment margin is unlikely to be an important driver of our results.

4.3. What prevents effective job search among refugees?

4.3.1. Quantitative evidence on liquidity constraints. What have we learned about barriers to job search among refugees? Our ESO follow-up surveys at the two-month and four-month marks show that the cash intervention has the largest impacts on job search and employment for this population. Here, we present evidence suggesting that the cash grant is effective because liquidity constraints are a key labour market barrier for refugees.[29] We do this by studying whether proxies for available liquidity are associated with higher control job search intensity and whether treatment effects are heterogeneous with respect to liquidity. Further, we report impacts on additional measures of job quality. Liquidity-constrained individuals would forgo desirable employment opportunities due to the inability to pay for search and application costs; if these constraints are binding, the marginal jobs obtained by cash beneficiaries would have similar or better quality as the control jobs.

First, among control refugees, we find a strong association between job-search intensity and expenditure at baseline (a proxy of liquidity). We plot this relationship non-parametrically in Figure A.10 in the Online Appendix; both the probability of searching for work and the number of job applications increase with expenditure, especially for individuals with expenditure below the median. These associations are sizeable. Using a linear regression, we find that a one standard deviation increase in expenditure at baseline is associated with a 0.6 standard deviation increase in the number of job applications sent (and a 0.08 standard deviation increase in the probability of job search). In contrast, among Jordanians, this relationship is much weaker: an increase in expenditure by one standard deviation is associated with a 0.3 standard deviation increase in the number of job applications and a 0.03 standard deviation increase in the probability of job search (also see Online Appendix Figure A.11). This is consistent with several recent quantitative reports from international organisations working with refugees, which report that cash-constraints are a first-order concern both for Syrian refugees in Jordan (Abu Hamad et al., 2017), and for Syrian refugees in Lebanon (Lehmann and Masterson, 2014; Chaaban et al., 2020); see Schuettler and Caron (2020) for a review. It is also consistent with our qualitative

---

[28] REACH (2021) collected data on 194 microenterprises ran by Syrian refugees in Jordan (with respondents who are broadly similar to the respondents in our project). In that sample, the average reported start-up requirement for a business was 697 JOD. Our transfer was less than 10% of that capital requirement: just 65 JOD.

[29] The analysis in this section was not pre-registered, but is part of our exploration of treatment mechanisms.

interviews – described shortly – in which we found that financial constraints are a key barrier, and that the cash grant was seen as facilitating job search.[30]

Second, we show that the impacts of the cash intervention are driven by refugees who have expenditure below the median. We show this in Online Appendix Figures A.12, A.13 and A.14: impacts on job search and employment are concentrated among the poorest respondents (while impacts on job applications are more evenly distributed). In contrast, the information and nudge intervention have (i) generally weaker impacts on job search among low-expenditure refugees, and (ii) employment impacts for low expenditure refugees that are about half of those of the cash intervention. Additional evidence in support of credit constraints comes from refugees' reports on how they spent the cash: 26 percent of recipients in the low-expenditure group report that they mostly spend the money on job search. Among above-median expenditure recipients, this proportion drops to 18 percent.[31]

Third, we find that the cash intervention boosts job retention and, after four months, hourly wages; we show this in Table 6. The grant doubles the probability of having retained a job between the two- and the four-month interviews – from 3.3 percent to 6.2 percent. Further, mean hourly wages among employed cash beneficiaries are 0.63 of a standard deviation higher than in the control group. The other two interventions, on the other hand, are associated with much smaller increases in retention and hourly wages (for example, job retention among information recipients is 5.2 percent, and the impact on hourly wages is about 0.04 of a standard deviation). We note that these wage effects need to be interpeted with some caution, given the positive employment effect; nonetheless, we note that the positive impact is large, and robust to a reasonable bounding exercise.[32] These impacts indicate that, consistent with the prediction of a model where job search is constrained by limited liquidity, the cash intervention enables jobseekers to find jobs that have higher match quality – and hence are more stable and better paid.

4.3.2. Eliciting expert forecasts. At the time of launching our interventions, we conducted an incentivized elicitation exercise with IRC staff. We ran this online, surveying 16 staff based in Amman, and four senior staff based in New York. For all respondents, we began by providing descriptive quantitative information on the

---

[30] Hagen-Zanker et al. (2018) also report on qualitative interviews with Syrian refugees in Jordan, and also emphasize the primary importance of financial constraints.

[31] Among Jordanians, 32 percent of respondents report to have spent the cash mostly on job search. However, this proportion does not vary by baseline expenditure. Given the null impacts on job search and the higher control job-search intensity, it is likely that cash given to Jordanians has mostly financed infra-marginal job search.

[32] Specifically, we have run an Attanasio et al. bounding exercise (Attanasio et al., 2011) for the result on hourly earnings at the four-month follow-up. The original coefficient is 0.3; we find that, if we bound using the 75th and 25th percentiles of the control distribution, we obtain bounds of 0.056 and 0.540 (respectively a 5 percent and a 46 percent increase). If we use the much more extreme approach of bounding using the 90th and 10th percentiles of the control distribution, we obtain bounds of -0.158 and 0.540 (respectively a 13 decrease and a 64 percent increase).

background of our sample, a brief description of each intervention, and information on the employment rate for a similar sample in 2018 (namely, a rate of 2.5%). We then asked a series of questions about each respondent's prediction for the rate of employment after six weeks; this was directly incentivized.[33] We illustrate forecast employment rates in the Online Appendix (Figures A.16 and A.17).

We have three main findings. First, relative to our estimated treatment effects, local staff were very optimistic in their forecasts: at the median, they predicted employment rates of 20%, 10% and 9% for the cash, information and nudge interventions respectively (the median forecast for the control group was 2.25%). Senior staff had more accurate forecasts: medians of 7%, 5% and 4% (against a median prediction in the control group of 3%). Second, the dispersion in forecasts was very large, indicating substantial uncertainty about treatment impacts. Third, interestingly, both local and senior staff correctly anticipated that the cash intervention would be most effective.

4.3.3. Qualitative fieldwork.   Five months after the trial began, we conducted structured qualitative interviews in the form of focus group discussions with participants who have received one of the three search interventions. The purpose of the interviews was to build a deeper understanding of the job search process and the mechanisms by which some interventions may have worked for different groups of participants. Participants were divided into six single-gender groups, each group focusing on one of the three interventions (cash, informational, or psychological support). We found the following results.

First, consistent with our experimental findings, respondents identified financial constraints as a key barrier to economic opportunity. The cash grant was by far the most popular intervention, irrespective of the demographics of the respondents. Among those who received the grant, some indicated that they used the cash directly to cover transportation costs when searching for work. Several others, in particular, Syrian women, reported using the cash to cover immediate basic needs such as medical bills for themselves or the family. Given participants' highly vulnerable economic situations, the cash was seen as necessary step before searching for work. One Syrian woman in the cash intervention reported that she used the cash for medical care, which allowed her to then begin searching for work and eventually to find a job in a factory. These qualitative findings match our quantitative results.

Second, participants reported mixed views on the informational and nudge interventions. The information intervention was only reported to be useful or accessible by Jordanians who had higher education levels than the typical participant in our study. At the same time, for some of those who understood the content, the information intervention was a source of frustration given that employers frequently

---

[33] Specifically, we told respondents that we would randomly draw one of their employment forecasts; if this forecast was within 1 percentage point of the correct answer, we would provide a ticket for a lottery with 50 tickets and a prize of US$200.

violated labor laws. As such, the information on labor laws was not deemed to represent de facto labor rights.

The nudge intervention was reported to be useful by the Syrian women we talked to, but for reasons quite different from our original theory of change. Rather than working as a commitment device, these respondents reported that the intervention motivated them by making them feel 'like someone cared about them', and by making job search top of mind. One Syrian woman indicated that following the last SMS message, she continued to use her tool and shared it with her female neighbours. However, others (mostly Jordanian males) reported a lack of interest or desire to receive SMS reminders on job search intentions and achievements. For this group of participants, the intervention was perceived not to be useful given they were already searching for work.

Finally, participants in all focus groups stressed that formal jobs are often poorly paid and offer bad working conditions. Women reported concerns about working in the formal manufacturing sector, where they would be required to mix with men and undergo long hours away from their children. In fact, many women indicated a strong interest in starting home-based businesses like a kitchen, beauty salon, accessory shop, or nursery which would offer them both a 'safer' and more flexible work environment. Men also perceived formal work in factories to be undesirable, but for different reasons. They mostly complained about low pay: wages in manufacturing firms are considered insufficient to support a family. Per diem informal work, mainly in services like plumbing, carpentry, and painting, was seen as more profitable. Additionally, factory work was seen as being potentially exploitative and men shared stories about wages being withheld or delayed by employers for arbitrary reasons. These observations suggest that constraints to labor market participation among both refugees and Jordanians are by no means restricted to the search frictions we have focused on in this study.
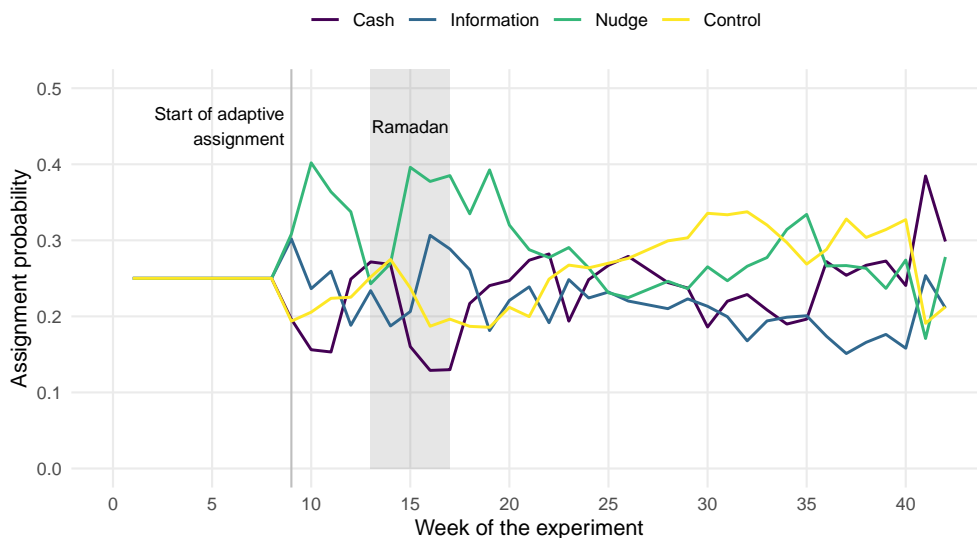
5. Assessment of performance of the Tempered Thompson Algorithm

We now consider various aspects of the performance of the Tempered Thompson Algorithm in our setting.

5.1. *Treatment assignment probabilities over time*

Consistent with the results presented in section 4.1, we find that in the last week of the study our algorithm places similar proportions of people in each of the four experimental groups. We show the probability of assignment to the four experimental conditions for each week of the study in Figure 1. By design, individuals are assigned to the different groups in equal proportion up to the sixth week of the study, as we have no information to update the priors up to that point. When learning started, the algorithm initially assigned more weight to the nudge intervention. However, this was slowly reversed after the 20th week of the study.

FIGURE 1. **Assignment probabilities by week**

The algorithm's departure from equal-proportions randomisation is somewhat more pronounced for specific strata. We show this in Figure E.A.1 in the Appendix, where we show strata-specific weekly treatment assignment probabilities, and in Table A.11 in the Online Appendix, where we show, for each treatment, the posterior probability that employment rates are highest under that treatment – that is, the posteriors that determine treatment assignment probabilities in our algorithm. While for some strata the assignment probabilities never depart from 25% in a sustained way, in some strata we do observe clear changes. For example, in the last week of the experiment, we assign almost 60% of inexperienced and less educated Syrian women to the cash intervention. Similarly, for some strata, the probability that the control is optimal drops to a few percentage points (e.g. inexperienced, less educated Syrian women). However, it should be stressed that, as discussed above, the differences in potential outcomes we estimate are small and hence the impacts of departing from equal-proportions randomizaton are limited in this context.

5.2. *Welfare contrasts*

Denote by $\hat{\theta}^{dt}$ the posterior expectation of θ^{dt} given all our data, at the end of the experiment. We now present three 'welfare contrasts' that quantify the overall impact of our interventions, both against a counterfactual where no treatment is given, and against a counterfactual where treatments are randomized in equal proportion. First, within the experiment, we compare the average potential outcomes for the *actually chosen treatment* assignment to the average that would have obtained under *random*

assignment,

$$\Delta_1 = \frac{1}{N} \sum_{i,t} \left(\hat{\theta}^{D_{it} X_{it}} - \frac{1}{4} \sum_d \hat{\theta}^{d X_{it}} \right).$$

This measures how much better we did for our experimental participants, compared to a conventional design with fully random assignment.

Second, we compare the *optimal targeted policy*, and the *optimal non-targeted policy*, to the default of *no intervention* (treatment 0),

$$\Delta_2 = \sum_x \left(\max_d \hat{\theta}^{dx} - \hat{\theta}^{0x} \right) \cdot p^x,$$

$$\Delta_3 = \max_d \sum_x \left(\hat{\theta}^{dx} - \hat{\theta}^{0x} \right) \cdot p^x.$$

The definition of Δ_2 allows the optimized d to depend on x, while the definition of Δ_3 requires the same d to be implemented for all x.

We estimate that overall impacts on employment after six weeks are small; Table 7 reports our corresponding estimates of the three welfare contrasts specified above. We have two key findings. First, if we compare the optimal targeted policy to a counterfactual where no intervention is given (welfare contrast Δ_2), we estimate a gain in employment of 1.7 percentage points (95% credible set: [0.001, 0.034]). Relative to the employment rate in the control groups, this amounts to a 35% increase in employment. The optimal non-targeted policy, on the other hand, delivers a gain in employment of about half of a percentage point (welfare contrast Δ_3), with a credible sets that includes zero (95% credible set: [-0.015, 0.27]). The difference in employment gains between these measures suggests that there may be some modest gains from targeting (though it should be noted that the confidence sets overlap). Overall, the percentage point effects are on the lower end of the impacts of ALMPs on employment reported in McKenzie (2017) (which are typically measured over a longer time frame).

In our study, adaptive randomization did not generate any six-weeks employment gains over standard randomization. We show this by reporting welfare contrast Δ_1, in Table 7, which is very close to zero.

5.3. *The impact of the Tempered Thompson Algorithm with the two-month employment objective*

We operationalized our Tempered Thompson Algorithm using the rapid, six-week follow-up surveys—where we found no significant impact of any treatment. What if, instead of using employment at the six-week rapid follow-up survey, we had instead targeted the two-month employment rate—where we found large impacts of cash, information and nudge treatments? To answer this question, we run a counter-factual simulation exercise. Specifically, we mimic the arrival rate of respondents (and their strata), but simulate employment outcomes using the observed two-month treatment effects (i.e., posterior means measured at the stratum level); we then

simulate alternative treatment assignments based on those simulated employment outcomes.

Figure 3a shows the main result from this exercise: it plots the average simulated share assigned, over time, to the best treatment (where 'best treatment' is allowed here to vary with the strata).[34] The figure confirms that the Tempered Thompson algorithm is quite capable of generating substantial learning in this kind of context, with the kinds of effect sizes that one might observe: indeed, the figure confirms that the limited performance of the algorithm in our experiment is attributable to our decision to target the six-week 'rapid follow-up' data, rather than the two-month ESO data.

In Figure 3b, we show the simulated expected employment rates by week. We estimate that the rate of employment in the control group would have been 11.2%, the rate of employment among respondents in a 'pure RCT' (with equal weight on each treatment) would be 12.4%, and the rate of employment from the Tempered Thompson algorithm would be 13.4%. That is, had we used the two-month data from the ESO surveys, rather than the six-week rapid follow-up data, Tempered Thompson would have generated approximately double the employment gains of a standard experiment.[35]

6. Conclusion

Randomized controlled trials have come under criticism from an ethical perspective. For example, Deaton (2020, p. 21) points out that "It is particularly worrying if the research addresses questions in economics that appear to have no potential benefit for the subjects." Relatedly, implementation partners might be reluctant to engage in continued experimentation if they believe that they already know which intervention works best. Adaptive experimentation can help mitigate both ethical criticisms of RCTs and the reluctance of implementation partners to engage in experimentation, by setting welfare maximization (or regret minimization) as the main objective of an experiment. Indeed, any optimal adaptive experimental design has the property that participant welfare cannot be increased in the long-run by using any other (experimental or non-experimental) treatment allocation procedure.
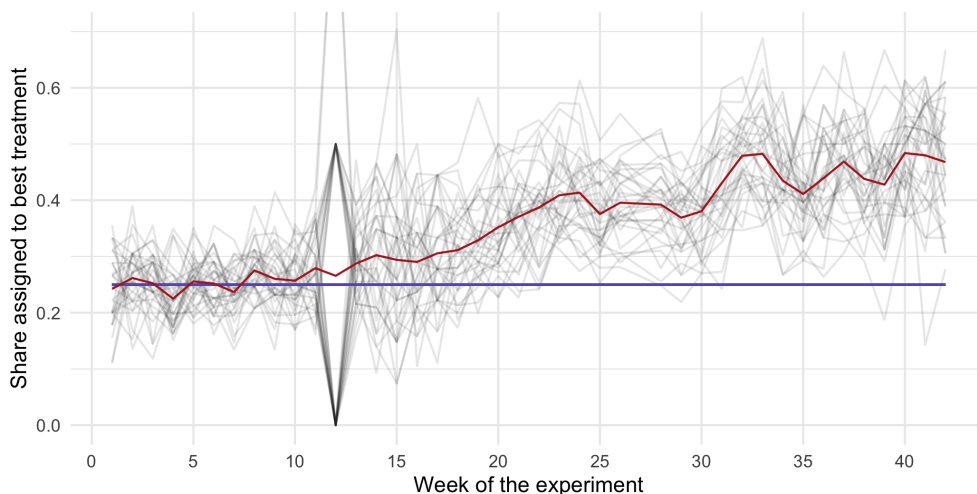
In this paper, we have reported the results of an implementation of adaptive targeted treatment allocation in a field experiment. Our Tempered Thompson Algorithm strikes a balance between maximizing participant welfare and providing precise estimates of treatment effects. Our implementation context was novel: We

---

[34] The volatility in this figure – and in Figure 3b – is driven by day-to-day changes in the proportions of different strata; this is reflected in the figures because our simulation replicated the exact arrival patterns of strata in our experiment. In week 12 there are only two observations in the underlying data, which explains the large volatility of assignment shares for this week.

[35] Appendix Figure A.15 shows the simulated gains by strata; the effect is particularly driven by Syrians with employment experience (both men and women, with and without high-school education), and by Jordanian men with high-school education and work experience.
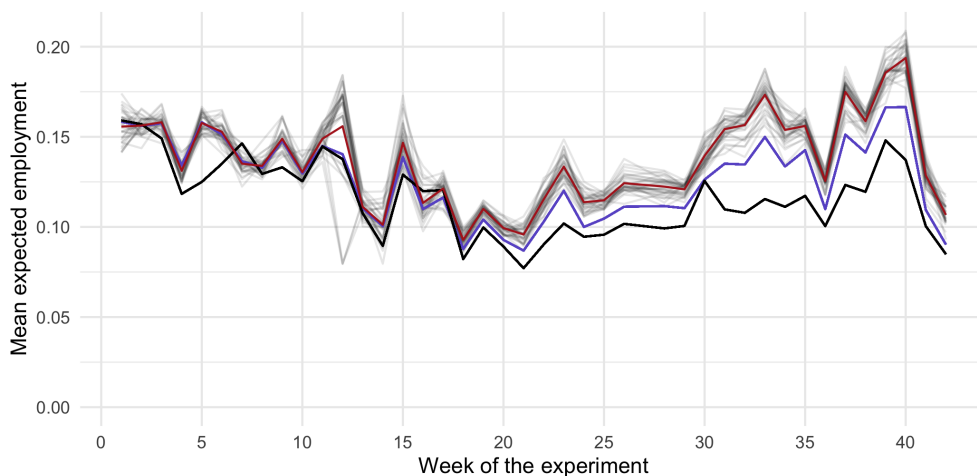
FIGURE 2. **Simulated performance of Tempered Thompson Algorithm**

(A) **Simulated probability of being assigned to the best treatment**

This plot shows simulated shares assigned to the best treatment within each stratum from 32 simulations (red),
contrasted with the corresponding shares from a pure RCT (blue).
The grey lines show each of the 32 simulation trajectories.

(B) **Simulated expected employment rate by week**

This plot shows simulated expected employment from 32 simulations (red),
contrasted with expected employment from a pure RCT (blue),
and the counterfactual where everybody receives the control treatment (black).
The grey lines show each of the 32 simulation trajectories.

looked at the effects of active labor market policies on Syrian refugees and local job-seekers in Jordan. Our treatments did not have a significant effect on refugee employment after six weeks, but the cash grant had a substantial impact on longer-term employment outcomes.

Our results show that adaptive targeted experiments can be straightforwardly deployed in the field and can be used to draw scientific and policy conclusions.

Moreover, our methodology creates many possibilities for further applications. The Tempered Thompson Algorithm is a powerful tool for any setting in which subjects arrive over time and their outcomes are observed within a short time-frame. In addition to employment programs, our methodology may be applied in many other development contexts, including drug and vaccination programs, agricultural technology adoption programs, climate change mitigation interventions, and emergency relief programs. Further work on adaptive experiments can tailor experimental designs to specific applications.

One of the most important considerations in future adaptive experiments will be defining the objective Y_t.[36] The choice of the objective can change the performance of the algorithm affecting short-run and long-run participant outcomes. In our case, the choice of target outcome, i.e., formal employment six weeks after the intervention, was mainly driven by the organizational objectives of our implementation partner and their donors. Our results in Section 5.3 suggest that, had we used a broader definition of employment, our adaptive algorithm would have quickly found the optimal intervention. More generally, in our context, as in many others (such as education and health), policymakers do not only care about short-run outcomes that adaptive treatment policies might typically target. But adapting treatment allocation based on long-term outcomes can often make the field experiment too long and costly. Instead of only measuring and targeting long-term outcomes, the designer might therefore wish to find a set of short-run proxies, i.e., 'statistical surrogates', for long-term welfare (Athey et al., 2019); an adaptive targeted field experiment would therefore be designed in order to target these statistical surrogates.[37]

---

[36] The challenge of aligning decisions based on machine learning (e.g., of robots and other artificial intelligence systems) with broader societal interests is by no means unique to adaptive experiments (Taylor et al., 2016; Russell, 2019).

[37] We discuss additional recommendations for future adaptive field experiments in Appendix B.

TABLE 4. **Treatment effects on main outcomes after 2 months**

	Employed	Earnings	Well-being	Social integration	Intends to migrate
Syrians					
Cash	0.052 (0.017)	7.204 (0.062)	0.021 (0.333)	-0.009 (0.575)	-0.038 (0.838)
Information	0.047 (0.036)	6.209 (0.092)	0.025 (0.309)	-0.035 (0.728)	-0.041 (0.856)
Nudge	0.035 (0.081)	4.210 (0.185)	0.007 (0.445)	-0.055 (0.817)	-0.031 (0.783)
Control mean	0.091	16.268	0.088	0.011	0.664
Observations	1608	1605	1608	1608	1598
Jordanians					
Cash	-0.007 (0.618)	-1.491 (0.618)	0.101 (0.012)		
Information	-0.006 (0.624)	-2.486 (0.696)	0.019 (0.342)		
Nudge	-0.004 (0.585)	-1.684 (0.631)	0.015 (0.385)		
Control mean	0.128	29.22	0.069		
Observations	1985	1977	1985		

TABLE 5. **Treatment effects on main outcomes after 4 months**

	Employed	Earnings	Well-being	Social integration	Intends to migrate
Syrians					
Cash	0.038 (0.027)	6.550 (0.040)	0.043 (0.163)	0.005 (0.472)	-0.044 (0.875)
Information	0.019 (0.148)	4.567 (0.105)	0.003 (0.480)	0.001 (0.470)	-0.046 (0.884)
Nudge	0.003 (0.449)	0.260 (0.484)	0.052 (0.106)	0.005 (0.467)	-0.034 (0.810)
Control mean	0.052	9.76	0.008	-0.005	0.675
Observations	1565	1563	1565	1565	1561
Jordanians					
Cash	-0.025 (0.855)	-7.515 (0.911)	0.068 (0.055)		
Information	-0.009 (0.658)	-4.299 (0.773)	0.041 (0.169)		
Nudge	-0.002 (0.544)	-2.845 (0.709)	0.040 (0.176)		
Control mean	0.144	33.451	0.039		
Observations	1913	1900	1913		

Note: These tables report treatment effects on the five main outcomes specified in the Pre-Analysis Plan, 2 and 4 months after the baseline interview. 'Employed' is a dummy for whether the person has a wage-paying job at the time of the interview. 'Earnings' is the value earnings from the main job (where individuals who are not in wage employment are assigned a zero). 'Well-being' is a weighted index that includes: (i) a measure of expenditure, (ii) a measure of positive affect, and (iii) a measure of life satisfaction. 'Social integration' is an index of social integration. 'Intends to migrate' is a dummy for whether the respondent intends to migrate to a third country (i.e. this measure does not include return migration). The first panel reports impacts for Syrians, and the second panel reports impacts for Jordanians. Next to each treatment effect estimate, we report a randomization inference p-value.

TABLE 6. **Retention and wages for Syrians**

	Job retention month 4	Hourly wage month 2	Hourly wage month 4
Cash	0.030 (0.036)	0.037 (0.393)	0.308 (0.055)
Information	0.017 (0.139)	-0.131 (0.835)	-0.087 (0.673)
Nudge	0.011 (0.270)	-0.124 (0.812)	-0.010 (0.527)
Control mean	0.034	1.377	1.197
Observations	1565	193	94

Note: This table reports treatment effects on the three additional outcomes. 'Job retention month 4? is a dummy capturing whether, at the time of the 4 month follow-up, the respondent holds the same job they held at the time of the 2 month follow-up. 'Hourly wage' is a variable obtained by diving monthly wage earnings by hours worked. Next to each treatment effect estimate, we report a randomization inference p-value.

TABLE 7. **Welfare contrasts**

	Estimate	95% Credible set
Δ_1	0.002	(0.000, 0.004)
Δ_2	0.017	(0.001, 0.034)
Δ_3	0.006	(-0.015, 0.027)

Note: The table reports the welfare contrasts defined in Section 5.2.

Appendix A: Tempered Thompson Algorithm as an approximate solution to a decision problem

In Section 3, we introduced the Tempered Thompson Algorithm and provided an asymptotic characterization of its behavior. We argued that the Tempered Thompson Algorithm has desirable properties and sensibly trades off several objectives, including inference and power (for treatment effects corresponding to all pairwise comparisons of treatments), and participants' welfare (as captured by their employment prospects). This approach, where an algorithm is proposed and then its properties are studied, corresponds to the approach taken in much of the bandit literature; see e.g. Bubeck and Cesa-Bianchi (2012); Lattimore and Szepesvári (2020).

An alternative approach, more in line with traditional economic theory, would fully specify a decision problem with objective function and constraints, and then aim to solve for the optimum. This is in general not possible for bandit-type problems, however, where optimal solutions are out of reach except in very simple cases; e.g. Weber et al. (1992).

What we *can* do is to specify an objective and constraints, and solve for an (infeasible) oracle solution to the corresponding decision problem. This oracle solution corresponds to *non-adaptive, non-random* allocation of the full sample to the different treatments. We then show that Theorem 1 implies that the *adaptive, random* allocation of the Tempered Thompson Algorithm asymptotically achieves the same performance as this oracle solution. An additional benefit of this approach is that it allows us to derive the optimal tuning parameter γ from more primitive parameters of the constraint and objective. We do so in the following discussion, which is then summarized in Propositions A.1 and A.2.

In conventional RCTs, a common objective is to minimize sample size, subject to the constraint of allowing for sufficiently precise inference on the effect of different treatments. This decision problem rationalizes the standard practice of stratification and power calculations for the determination of the required sample size. For conventional bandit algorithms, the objective is to maximize average participant outcomes for a given sample size. The solution of this problems requires quickly shifting assignment toward the best-performing treatments.

Our experiment has the objective of maximizing participant outcomes (i.e., employment prospects), as in bandits, subject to the constraint of allowing for sufficiently precise inference on the effect of different treatments, as in conventional RCTs.

More concretely, we require *frequentist (i.e., ex-ante, worst-case) guarantees* on the precision of estimates of the average structural function, which hold for any possible parameter vector θ_0. This requirement of frequentist guarantees implies constraints on the allowed allocation, and a lower bound on γ. The *oracle-optimal allocation* maximizes participant outcomes given knowledge of θ_0, subject to the frequentist constraints. This oracle solution allocates as many participants as possible to the best treatment arm in each stratum, subject to allocating enough participants to

each treatment in each stratum. Since in practice we don't know $\boldsymbol{\theta}_0$, the oracle-optimal allocation cannot be implemented. Instead, we implement an adaptive algorithm which converges to the oracle-optimal allocation.

A.1. Decision problem

The constraint of frequentist inference. Throughout our discussion in this section, we assume that the number N^{dx} of observations assigned to treatment d in stratum x is non-random; we consider a fixed allocation of the sample. After completion of the experiment, we estimate the average potential outcome for each treatment by $\hat{\theta}^d = \sum_x p^x \bar{Y}^{dx}$.[38] For a non-random allocation across strata and treatments, the variance of this estimator is given by

$$\mathrm{Var}_{\boldsymbol{\theta}_0}\left(\hat{\theta}^d\right) = \mathrm{Var}_{\boldsymbol{\theta}_0}\left(\sum_x p^x \cdot \bar{Y}^{dx}\right) = \sum_x (p^x)^2 \cdot \mathrm{Var}_{\boldsymbol{\theta}_0}\left(\bar{Y}^{dx}\right).$$

Our frequentist (i.e., ex-ante, worst-case) constraint in the experiment is that

$$\max_x \max_{\boldsymbol{\theta}_0} p^x \, \mathrm{Var}_{\boldsymbol{\theta}_0}\left(\bar{Y}^{dx}\right) \leq \lambda, \tag{A.1}$$

for a pre-specified $\lambda > 0$. Note that we require this bound to hold for all strata x and parameter vectors θ_0, which allows us to perform inference on treatment effect heterogeneity. Constraint (A.1) implies a corresponding bound of λ on $\mathrm{Var}_{\boldsymbol{\theta}_0}\left(\hat{\theta}^d\right)$, regardless of the true parameter vector $\boldsymbol{\theta}_0$. This latter requirement is closely related to conventional power calculations, as we show in Appendix A.2.

Recall that we have a non-random number of observations N^{dx} for treatment d in stratum x, and that Y is binary. Now, Constraint (A.1) can be solved explicitly. Since

$$\max_{\boldsymbol{\theta}_0} \mathrm{Var}_{\boldsymbol{\theta}_0}\left(\bar{Y}^{dx}\right) = \max_{\theta_0^{dx}} \frac{\theta_0^{dx}(1 - \theta_0^{dx})}{N^{dx}} = \frac{1}{4N^{dx}},$$

we require that

$$\max_x \frac{p_x}{4N^{dx}} \leq \lambda,$$

so we get the constraint on the sample allocation that

$$\forall x: \quad N^{dx} \geq \frac{p_x}{4\lambda}.$$

Denote the share of observations in stratum x that is allocated to treatment d by $\bar{q}^{dx}$. Then, by definition, $N^{dx} = N \cdot p^x \cdot \bar{q}^{dx}$, and the constraint becomes

$$\bar{q}^{dx} \geq \frac{1}{4 \cdot N \cdot \lambda}. \tag{A.2}$$

---

[38] We omit the time subscript t throughout, since it is not required for our discussion.

In our application, we choose an upper bound on $p^x \cdot \text{Var}_{\boldsymbol{\theta}_0}\left(\bar{Y}^{dx}\right)$ of $\lambda = .00125$. As we show in subsection A.2 below, this choice of λ is equivalent to a worst-case bound on the standard error of average treatment effect estimators of .05, or to a worst-case power of .8 for treatment effects of size .124. Our choice of λ implies a maximum variance of $\hat{\theta}^d$ equal to

$$\text{Var}_{\boldsymbol{\theta}_0}(\hat{\theta}^d) = \sum_x p^x \left[p^x \cdot \text{Var}_{\boldsymbol{\theta}_0}\left(\bar{Y}^{dx}\right)\right] \le \sum_x p^x \lambda = \lambda = .00125.$$

Note that for more realistic $\boldsymbol{\theta}_0$, the variance of these estimators will generally be much smaller. For an expected number of participants of $N = 4000$ (we ended up with slightly fewer in the actual experiment) our choice of λ implies that $N \cdot \lambda = 5$, and hence

$$\bar{q}^{dx} \ge .05. \tag{A.3}$$

The objective of participant welfare. Conventional RCTs are rationalized by minimizing the total sample size subject to the constraint (A.1). Solving this problem yields the standard recommendation of stratified treatment assignment, and the choice of sample size by power calculations; cf. Athey and Imbens (2017).

Rather than minimizing sample size, our objective is to choose the assignment share $\bar{q}^{dx}$ to maximize average expected participant outcomes

$$\sum_{d,x} p^x \cdot \bar{q}^{dx} \cdot \theta_0^{dx}. \tag{A.4}$$

The infeasible oracle solution. Suppose for a moment that $\boldsymbol{\theta}_0$ were known, and that we did not face the frequentist constraint (A.1). The oracle optimal solution to this unconstrained decision problem of maximizing (A.4) would set $\bar{q}^{dx} = 1$ if $d = d^{*x} = \arg\max_{d'} \theta^{d'x}$, and $\bar{q}^{dx} = 0$ otherwise. Imposing the frequentist constraint (A.1), but still assuming that we knew $\boldsymbol{\theta}_0$, the oracle optimal solution becomes

$$\bar{q}^{dx} = \begin{cases} 1 - (k-1)\frac{1}{N \cdot \lambda} & \text{if } d = d^{*x} \\ \frac{1}{N \cdot \lambda} & \text{else.} \end{cases} \tag{A.5}$$

For our application, where $N \cdot \lambda = 5$ and $k = 4$, we obtain

$$\bar{q}^{dx} = \begin{cases} .85 & \text{if } d = d^{*x} \\ .05 & \text{else.} \end{cases}$$

Tempered Thompson as a feasible algorithm. The "oracle-optimal" allocation described by Equation (A.5) is not feasible, since it depends on knowledge of d^{*x} for all x. Feasible algorithms can, however, approximate this choice of assignment shares $\bar{q}^{dx}$ in a data-dependent, adaptive manner. Our core result, Theorem 1, shows, in particular, that in large samples the Tempered Thompson Algorithm is as good as the oracle optimal allocation. In particular, item 2 of Theorem 1 states that the cumulative share q_t^{dx} allocated to treatment d in stratum x converges in probability to $\bar{q}^{dx}$ as defined in Theorem 1.

Note in particular that γ, the tuning parameter in the definition of the Tempered Thompson Algorithm, as in Assumption 1, is now pinned down by the parameters of the decision problem. By Equation (A.2), we need to set $\gamma = \frac{1}{4 \cdot N \cdot \lambda}$, or, for our application, $\gamma = .05$. This, of course, is exactly the tuning parameter used in our application.

We summarize the preceding discussion in the following proposition.

PROPOSITION A.1. *Under Assumption 1, consider the problem of choosing assignment shares $\bar{q}^{dx}$ to maximize participant welfare (A.4) (for a given $\boldsymbol{\theta}_0$) subject to the frequentist inference constraint (A.1) on the worst-case estimator variance. Then the following statements hold true:*

1. *The (infeasible) oracle solution to this problem is given by*

$$\bar{q}^{dx} = \begin{cases} 1 - (k-1)\gamma & \text{if } d = d^{*x}, \\ \gamma & \text{else}, \end{cases}$$

   where $\gamma = \frac{1}{N \cdot \lambda}$.
2. *If treatment is assigned using the Tempered Thompson Algorithm, then the cumulative share q_t^{dx} allocated to treatment d in stratum x converges in probability to $\bar{q}^{dx}$.*

A.2. Connecting the frequentist constraint to treatment effects and power calculations

We now discuss two alternative criteria which are implied by the frequentist inference constraint. We have described the constraint in terms of an upper bound on the variance of estimators of the average structural function (i.e., average potential outcomes). Two alternative constraints are (i) in terms of the variance of estimators for the average treatment effect for each pair of treatment contrasts (i.e., standard error), and (ii) in terms of the power of statistical tests.

Constraint on standard error bounds. For each pair d, d' of treatment values, consider the following estimator of the average treatment effect comparing d and d',

$$\hat{\Delta}^{d,d'} := \sum_x p^x \bar{Y}_t^{dx} - \sum_x p^x \bar{Y}_t^{d'x}.$$

Assuming again a non-random allocation across strata and treatments, the variance of this estimator is given by

$$\operatorname{Var}_{\boldsymbol{\theta}_0}\left(\hat{\Delta}^{d,d'}\right) = \sum_x (p^x)^2 \cdot \left[\operatorname{Var}_{\boldsymbol{\theta}_0}\left(\bar{Y}_t^{dx}\right) + \operatorname{Var}_{\boldsymbol{\theta}_0}\left(\bar{Y}_t^{d'x}\right)\right].$$

If $\max_x p^x \cdot \operatorname{Var}_{\boldsymbol{\theta}_0}\left(\bar{Y}_t^{dx}\right) \le \lambda$, as we required above, then

$$\max_{d,d',x} p^x \cdot \left[\operatorname{Var}_{\boldsymbol{\theta}_0}\left(\bar{Y}_t^{dx}\right) + \operatorname{Var}_{\boldsymbol{\theta}_0}\left(\bar{Y}_t^{d'x}\right)\right] \le 2\lambda.$$

Choosing $\lambda = .00125$, as as above, we get an upper bound on the standard error for these estimated average treatment effects of $\sqrt{2 \cdot .00125} = .05$.

Constraint on power of statistical tests. Bounds on the variance (standard error) of $\hat{\Delta}^{d,d'}$ and power calculations are equivalent. To see this, consider an assumed true treatment effect of $\Delta^{d,d'}$, and a required power of β, such as $\beta = .8$. Under the usual asymptotic normal approximation, and for a one-sided test of level .05 for the null of no effect, this lower bound on power can be written as

$$P\left(\frac{\hat{\Delta}^{d,d'}}{\sqrt{\text{Var}_{\boldsymbol{\theta}_0}(\hat{\Delta}^{d,d'})}} > 1.64\right) = \Phi\left(\frac{\Delta^{d,d'}}{\sqrt{\text{Var}_{\boldsymbol{\theta}_0}(\hat{\Delta}^{d,d'})}} - 1.64\right) \geq .8.$$

This bound is in turn equivalent to requiring

$$\text{Var}_{\boldsymbol{\theta}_0}(\hat{\Delta}^{d,d'}) \leq \left(\frac{\Delta^{d,d'}}{\Phi^{-1}(.8) + 1.64}\right)^2.$$

For $\Delta^{d,d'} = .05 \cdot (\Phi^{-1}(.8) + 1.64) = .124$, we obtain the constraint $\text{Var}_{\boldsymbol{\theta}_0}\left(\hat{\Delta}^{d,d'}\right) \leq .0025$.

We summarize again in the following proposition, which helps guide the choice of γ for the Tempered Thompson Algorithm:

PROPOSITION A.2 (Equivalent constraints).
 Under Assumption 1, for non-adaptive assignment with fixed assignment shares $\bar{q}^{dx}$, the following two constraints are equivalent:

(1) *Constraint on the worst-case estimator variance:*
    $\max_x \max_{\boldsymbol{\theta}_0} p^x \, \text{Var}_{\boldsymbol{\theta}_0}\left(\bar{Y}^{dx}\right) \leq \lambda.$
(2) *Constraint on the minimal assignment share:*
    $\bar{q}^{dx} \geq \gamma$, *where* $\gamma = \frac{1}{4 \cdot N \cdot \lambda}$.

Constraints (1) and (2) imply the following two constraints:

(3) *Constraint on the worst-case standard error of average treatment effect estimators:*
    $\sqrt{\text{Var}_{\boldsymbol{\theta}_0}\left(\hat{\Delta}^{d,d'}\right)} \leq \sqrt{2\lambda}.$
(4) *Constraint on the power of statistical tests:*
    $P\left(\frac{\hat{\Delta}^{d,d'}}{\sqrt{\text{Var}_{\boldsymbol{\theta}_0}(\hat{\Delta}^{d,d'})}} > 1.64\right) \geq \beta$, *for* $\Delta^{d,d'} = \sqrt{2\lambda} \cdot (\Phi^{-1}(\beta) + 1.64).$

Appendix B: Lessons and recommendations for future adaptive field experiments

B.1. Short term employment as a surrogate

In our application, we designed the algorithm to maximize the short-term employment outcome of participants. As we discussed in Section 5.3, we would have been more successful had we been able to target longer-term employment.

We might therefore ask whether using short-term employment as a proxy-outcome is systematically misleading—or simply less informative—about long-term welfare. This is the question discussed in the literature on "surrogate outcomes"(see, e.g., Athey et al. (2019) for a recent review). Suppose that we are interested in the long-term outcome $\tilde{Y}$, and that short term employment Y satisfies the so-called "surrogacy condition",

$$\tilde{Y} \perp D | X, Y.$$

This condition requires that the long-term outcome should be independent of the treatment assignment conditional on observable characteristics X and short-run outcomes. Then, by the law of iterated expectations, we get that

$$E[\tilde{Y}|D, X] = E[[\tilde{Y}|D, X, Y]|D, X] = E[[\tilde{Y}|X, Y]|D, X].$$

An immediate implication of this result is that an algorithm maximizing the average of Y will also maximize the average of $\tilde{Y}$ if (i) the surrogacy condition holds, and (ii) $[\tilde{Y}|X, Y]$ is increasing in Y for all X.

Ex post, we can empirically check the surrogacy condition in our data, by regressing $\tilde{Y}$ on D, Y, X. The surrogacy condition is supported if we do not find a predictive effect of D on $\tilde{Y}$ given Y and X. We test the surrogacy condition for 4-month (long-term) employment outcome $\tilde{Y}$ using either 6-week or 2-month (short-term) employment outcomes Y as surrogates. The results are shown in Table B.A.1. For either subgroup (Syrians and Jordanians) and either short-run surrogate outcome Y (6-week employment and 2-month employment) we do not reject the null hypothesis of the surrogacy condition; cf. the p-values for the Wald test in Table B.A.1. That said, for 6-week employment it does appear that cash is marginally predictive of 4-month employment, suggesting a possible violation of the surrogacy condition. Overall, these findings suggest that using short-term employment as a proxy-outcome is not systematically misleading, though it is not very informative either.

B.2. Non-binary outcomes

In our discussion and empirical application we have focused on binary outcomes. This is motivated by the nature of our application, where the outcome of interest is employment status at the follow-up survey. Nothing prevents us, however, from considering non-binary outcomes for the Tempered Thompson Algorithm. For normally distributed outcomes, one could – for instance – consider the following

TABLE B.A.1. **Test of the surrogacy condition**

	Dependent variable: 4-month employment			
Cash	0.037	0.026	-0.026	-0.026
	(0.017)	(0.016)	(0.021)	(0.020)
Information	0.024	0.007	-0.000	0.000
	(0.019)	(0.018)	(0.022)	(0.021)
Nudge	-0.001	-0.007	-0.002	0.002
	(0.014)	(0.014)	(0.020)	(0.019)
6-week employment	0.358	0.257		
	(0.073)	(0.034)		
2-month employment			0.497	0.452
			(0.044)	(0.035)
Wald test (p)	0.081	0.245	0.575	0.527
Sample	Syrian	Jordanian	Syrian	Jordanian

Note: Robust standard errors in parentheses. All regressions are re-weighted using strata-specific weights. The second row from the bottom reports the p-value from a Wald test of the joint null hypothesis that the coefficients on all the treatment dummies are equal to zero.

hierarchical Bayesian model:

$$Y_{it}^d|(X_{it} = x, \theta^{dx}, \mu^d, \tau^{2d}, \sigma^2) \sim N(\theta^{dx}, \sigma^2),$$
$$\theta^{dx}|(\mu^d, \tau^{2d}, \sigma^2) \sim N(\mu^d, \tau^{2d}),$$
$$(\mu^d, \tau^{2d})|\sigma^2 \sim \pi_1,$$
$$\sigma^2 \sim \pi_2.$$

As before, such a model leads to an optimal combination of within and between stratum information, leading to posterior estimates of θ^{dx} which can be used in the Tempered Thompson Algorithm. More complicated outcome models – for example, allowing for mixture distributions, truncation, and so on – can be similarly defined and implemented: the Tempered Thompson Algorithm can be applied to any hierarchical Bayesian model that generates a posterior distribution of the mean of the potential outcomes Y_{it}^d, conditional on covariates.

B.3. *Choosing the sample size*

In conventional experiments, the sample size is often chosen based on power calculations. For a given effect size, the sample size is set such that the probability of rejecting the null hypothesis of no effect equals or exceeds a conventional value (e.g., 80%) for a test using a conventional significance level (e.g., 5%).

In our context, this approach does not apply directly, as our objective is not simply to satisfy a particular conventional power constraint. Instead, it would be more natural to choose the sample size according to cost-benefit calculations. Even better, one might use a data-dependent stopping rule to maximize welfare.

In our setting, increasing the sample size (number of waves and/or number of participants per wave) impacts (i) the welfare of participants and potential participants,

(ii) the precision of estimates (and hence the power of statistical tests), and (iii) the cost of implementing the experiment. Let us consider each of these in turn.

For small sample sizes, the (Tempered) Thompson Algorithm assigns roughly equal shares of observations to the different treatment arms. Correspondingly, the welfare of initial participants equals the average of the θ^{dx} across treatment arms d. As the algorithm learns the (conditionally) optimal treatment d^*, the welfare of additional participants increases, and then converges to the constrained optimum, as described by Part 3 of Theorem 1. Therefore, as a function of sample size, total participant welfare shows increasing returns to scale initially, and eventually becomes linear.

The same argument characterizes the total cost of the experiment. Per participant costs initially equal the average of costs across treatment arms, and then converge to

$$(1 - \gamma) \times (\text{cost of the best treatment}) + \gamma \times (\text{average cost across all treatments}).$$

Estimator precision and power, on the other hand, exhibit decreasing returns to scale. This is for two reasons. First, and this is true generally, the variance for each treatment arm is inversely proportional to the sample size for that treatment arm. Second, and this is specific to our algorithm, sample sizes for different treatment arms become more unbalanced over time, reducing the marginal decrease of variances for treatment effect estimation.

The following (myopic) rule provides a good rule-of-thumb for choosing when to stop the experiment, once we are in the regime of Theorem 1 (where we can, to a first order, ignore the dynamic returns of learning): In each period, calculate the expected return to and cost of an additional wave. We can do so based on the assignment shares $(1 - \gamma)\hat{p}_t^{dx} + \gamma/k$ implied by the algorithm, as well as the expected participant welfare θ^{dx} for each stratum and treatment arm as calculated by the hierarchical Bayesian model. The net return is given by the sum of the weighted average of expected participant welfare net of costs (averaged across treatment arms with weights given by the assignment shares) and the expected decrease of treatment effect estimator variance based on the expected increase of sample size for each arm. Continue the experiment if and only if this expected net return exceeds 0.

B.4. Choosing the wave size

There is a separate question that arises conditional on the total sample size: *How many waves should there be?* (Equivalently, *how large should each wave be?*) Theory provides guidance. In terms of large sample behavior, it turns out that the wave size does not matter.

Intuitively, larger waves reduce the short-run possibility of adaptation. A design with smaller waves can always replicate a design with larger waves by ignoring the most recent information. But the delay in learning induced by larger wave sizes becomes relatively negligible as the sample size increases. That said, in smaller samples, choosing smaller waves (given total sample size) leads to slightly better performance, because of the possibility to adapt assignment shares more quickly.

On the other hand, practical constraints might restrict the frequency of waves, based on the time it takes fo outcomes to be realized and measured, or based on institutional considerations (e.g., school children arriving in cohorts). Furthermore, there might be limits on the total run-time of the experiment. As a matter of practice, then, we recommend to choose waves as small as possible, subject to such practical constraints.

B.5. Choosing a prior

In the construction of the algorithm used in our application, we opted for a symmetric default prior with a hierarchical structure that implies prior independence across treatment arms. The use of such a prior reflects researcher equipoise relative to different alternatives. In particular, it implies that the initial phase of the experiment looks like a classic RCT with symmetric assignment probabilities across treatment arms. In future work, experimenters may however consider alternatives to our preferred approach.

First, they may consider incorporating information from expert forecasts in the prior. This will not affect the long-run behavior of the algorithm (unless the prior is degenerate), but it may influence its short-run performance. If expert beliefs are correct, using these beliefs will lead to higher welfare for initial participants. If they are wrong, using them could lead to lower welfare for initial participants, and slower learning of true effects. In our setting, the experts we interviewed were overly optimistic about the impacts of the interventions. The use of such overly optimistic priors would slow down convergence of the algorithm. We furthermore have limited information about expert uncertainty, which complicates the specification of a complete prior. In future work, it might be interesting to explore how to use expert forecasts to form a prior that can be incorporated in the algorithm.

Second, experimenters may consider relaxing the assumption of within-stratum independence across treatment arms. Again, this is an assumption that may affect the speed of convergence to the optimal allocation, but does not influence the long-run behavior of the algorithm. Relaxing this assumption would be particularly appropriate in settings with many different treatment arms where some groups appear to have systematically higher outcomes, regardless of treatment. The recent work by Kasy and Teytelboym (2023) provides examples of hierarchical priors based on two-way random effects models, such as the following, which allows for within-stratum dependence:

$$Y_{it}^d | (X_{it} = x, \theta^{dx}) \sim Ber(\theta^{dx}),$$
$$\theta^{dx} = \frac{1}{1 + \exp\left(-(\gamma^d + \delta^x + \varepsilon^{dx})\right)}$$
$$\gamma^d \sim N(0, \tau_d^2), \quad \delta^x \sim N(0, \tau_x^2), \quad \varepsilon^{dx} \sim N(\mu, \tau_{dx}^2).$$

B.6. *Alternative adaptive assignment algorithms*

The purpose of Tempered Thompson Algorithm proposed in this paper is to trade off participant welfare and estimator precision. There are, of course, many alternative algorithms which have been discussed in the multi-armed bandit literature (see, e.g., Slivkins (2019) for a review).

A particularly simple alternative would be to first run a traditional RCT, with equal assignment shares across treatment arms. After a certain number of observations, one could then switch to the "greedy" choice of assigning the best-performing treatment to all future participants. This is called the "Explore-First" algorithm. This procedure can, in large samples, achieve similar welfare as the Tempered Thompson Algorithm, since there is a limited need for exploration, and thus for balancing of exploration and exploitation.

For smaller samples, the Explore-first algorithm is dominated by procedures based on Thompson sampling (such as the Tempered Thompson Algorithm; see also the discussion in Section 1.2 and the remainder of Chapter 1 in Slivkins (2019)).

B.7. *Inference and the winner's curse*

We have discussed inference for various objects in Section 3.5. In particular, we considered both Bayesian inference and randomization inference for individual treatments, and for the expected welfare gains due to the Tempered Thompson Algorithm. Inference on expected welfare gains is related to the problem of "inference on winners" (Andrews et al., 2019) although it is orthogonal to the algorithm used for treatment assignment. Remarkably, neither Bayesian inference nor randomization inference need to explicitly correct for the selection issue of focusing on winners. For Bayesian inference, this holds because the posterior already conditions on the full data. The choice of "winner" is a function of the data. Conditioning on the identity of the winner, then, does not carry any further information which would need to be reflected in the posterior.

Randomization inference compares a statistic to the distribution of this same statistic over permutations (re-assignments) of treatment, according to the assignment algorithm used. No matter which statistic is used, this results in a test which controls size (the probability of rejection under the null) for the null hypothesis that all potential outcomes are the same. In particular, this is true for statistics such as the largest mean outcome across treatments. For any re-assignment of treatment, the largest mean might of course correspond to another treatment than in the actual data.

In this paper, we have not considered Neyman-style inference which needs to account for selection; this is akin to multiple testing corrections (Andrews et al., 2019). Implementing such inference remains an interesting avenue for future work.

B.8. Non-stationarity

In our hierarchical Bayesian model, we have assumed that the data-generating process is stationary, i.e, the distribution of potential outcomes, conditional on covariates, is constant over time. This assumption might be problematic in our setting. Potential outcome distributions change over time. However, this does not mean that our approach is flawed. In the multi-armed bandit literature, non-stationary settings are considered using the framework of "adversarial" bandits, cf. Cesa-Bianchi and Lugosi (2006). This branch of the literature provides worst-case performance guarantees for any sequence of potential outcomes, rather than just for sequences of potential outcomes generated by i.i.d. draws from a fixed distribution, as in our Theorem 1. Regret guarantees can be provided even for this adversarial setting, where regret is defined relative to the comparison point of any *fixed* assignment policy, mapping covariates into a distribution over treatments.

Remarkably, bandit algorithms (such as Thompson sampling) are able to achieve worst-case average regret (relative to any fixed treatment or targeted assignment policy) on the order of $1/\sqrt{T}$ (up to logarithmic terms). This is no worse than the performance for the worst-case stochastic setting, where potential outcomes are i.i.d. draws. Put differently, algorithms such as the Tempered Thompson Algorithm perform well (in terms of participant welfare) if there exists any fixed assignment policy that performs well—no matter how the outcomes are generated (see, e.g., Chapter 3 of Bubeck and Cesa-Bianchi 2012, and Chapter 6 of Slivkins 2019).

It is also instructive to consider the sequences of potential outcomes leading to worst-case behavior. Lewi et al. (2020) provide a detailed characterization for Thompson sampling for the bit-prediction setting. They show that for this setting the worst-case form of non-stationarity *delays learning for as long as possible*. Holding constant average potential outcomes across time, the worst-case sequence is obtained by shuffling outcomes across time such that in initial periods outcomes are balanced, and posteriors converge towards a precise belief that the true effect equals zero. Then, in the latter part of the time window, maximum mean differences imply that maximum regret is incurred due to the algorithm essentially sticking to the assignment probabilities of a static RCT, which are derived from the beliefs learned during the beginning of the time window.

B.9. Budget constraints

One final consideration is that of budget constraints. The existing literature considers treatment assignment under budget constraints both in the static setting (e.g., Bhattacharya and Dupas (2012)) and in the adaptive setting (e.g., Wang and Chen (2018); Kasy and Teytelboym (2023)). A natural way to introduce a budget constraint would be to have a constraint on the total number of units that can receive a particular treatment in a given round. Kasy and Teytelboym (2023) develop an adapted variant of the Tempered Thompson Algorithm that implements this feature in the "combinatorial

bandit" setting — a setting where the experimenter is interested to study a large number of possibile combinations of different treatments.

B.10. Does strata size matter?

In practical applications of the Tempered Thompson Algorithm, it might happen that strata are of different size, and that some strata will only have a limited number of observations. Does this affect the performance of the algorithm? The long-run behavior of the algorithm is not impacted by the size of different strata, and is characterized by Theorem 1. The short-run benefits from targeting hinge on having enough power for estimating treatment effect heterogeneity. Thus, small strata sizes will limit the benefits from targeting. The benefits from adaptivity may still be large though, as they hinge on having enough power for distinguishing *average* treatment effects from 0.

   To gain some intuition for how the hierarchical Bayesian approach deals with different strata sizes, consider two extreme cases. First, assume there are *no* prior observations in a given treatment/covariate cell. In that case, the conditional posterior mean for that cell is equal to the posterior mean for that treatment, across all strata, $\alpha^d/(\alpha^d + \beta^d)$. Put differently, the resulting treatment choice is the optimal choice without targeting. At the opposite extreme, if there are a lot of observations in a given treatment/covariate cell, then the posterior mean for this cell is simply the corresponding sample average, ignoring observations for all other covariate values. Finally, for intermediate cases, the posterior optimally interpolates between these two extremes. A key implication of this is that observations in small strata do not have disproportionate influence on their stratum-level posterior, nor on our estimates of the gains from targeting (if all strata were extremely small, we would mechanically estimate no gains from targeting.)

Appendix C: Proofs

C.1. Preliminaries

Our characterization of the large sample properties of our Tempered Thompson Algorithm relies on the following two useful results from the literature. The first is a law of large numbers for adaptive sequences, which can be found as Lemma 5 in Russo (2016). The second is a sufficient condition for consistency of Bayesian posteriors, known as Schwartz's theorem, which can be found as Theorem 6.16 in Ghosal and Van der Vaart (2017).

LEMMA C.C.1 (LLN for adaptive sequences). *Let $\{Y_n\}$ be an i.i.d sequence of real-valued random variables with finite variance and let $\{W_n\}$ be a sequence of binary random variables. Suppose each sequence is adapted to the filtration $\{\mathscr{H}_n\}$, and define $Z_n = P(W_n = 1|\mathscr{H}_{n-1})$. If, conditioned on $\mathscr{H}_{n-1}$, each Y_n is independent of W_n, then with probability 1,*

$$\lim_{n\to\infty} \sum_{l=1}^{n} Z_l = \infty \Rightarrow \lim_{n\to\infty} \frac{\sum_{l=1}^{n} W_l Y_l}{\sum_{l=1}^{n} Z_l} = E[Y_1].$$

THEOREM C.1 (Schwartz). *If $p_0 \in KL(\Pi)$ and for every neighborhood $\mathscr{U}$ of p_0 there exist tests φ_n such that $P_0^n \varphi_n \to 0$ and $\sup_{p\in\mathscr{U}_c} P^n(1 - \varphi_n) \to 0$, then the posterior distribution $\Pi(\cdot|X,\dots,X)$ in the model $X,\dots,X|p \sim^{iid} p$ and $p \sim \Pi$ is strongly consistent at p_0.*

In the statement of this theorem, Π is the prior distribution, $KL(\Pi)$ is its Kullback-Leibler support.

We furthermore leverage the following result from Melfi and Page (2000). This is a re-statement in our notation of their Corollary 2.1 (item 1 in the following theorem) and their Theorem 3.2, specialized to Bernoulli outcomes (item 2 in the following theorem).

THEOREM C.2 (Adaptive central limit theorem). *Suppose that Assumption 1 holds, except that we allow for general adaptive assignment algorithms determining D_{it}.*

1. *Suppose that D_{it} is a measurable function of X_{it}, the history $(D_{is}, Y_{is})_{s<t}$, as well as an independent randomization device. Let I_s^{dx} be the index pair (i,t) of the s-th time for which $D_{it} = d$ and $X_{it} = x$. Then the sequence $(Y_{I_1^{dx}}, Y_{I_2^{dx}}, \dots)$ is i.i.d. $Ber(\theta_0^{dx})$. These sequences are furthermore independent across different values of d, x.*
2. *Suppose additionally that the cumulative share q_t^{dx} allocated to treatment d in stratum x converges in probability to $\bar{q}^{dx}$ for all d, x, and the share of units in*

stratum x converges in probability to p^x. Then

$$\left(\sqrt{M_t \cdot \bar{q}^{dx} \cdot p^x} \, \frac{\bar{Y}_t^{dx} - \theta_0^{dx}}{\sqrt{\theta_0^{dx}(1 - \theta_0^{dx})}}, \sqrt{M_t \cdot \bar{q}^{d'x'} \cdot p^{x'}} \, \frac{\bar{Y}_t^{d'x'} - \theta_0^{d'x'}}{\sqrt{\theta_0^{d'x'}(1 - \theta_0^{d'x'})}} \right) \to^d (Z_{dx}, Z_{d'x'})$$

as $t \to \infty$, where $Z_{dx}, Z_{d'x'}$ are independent normal variables, for any pair $(d,x) \neq (d',x')$.

C.2. Proof of Theorem 1

Let $W_{it} = \mathbf{1}(D_{it} = d, X_{it} = x)$, and

$$Z_{it} = E_t[W_{it}] = \left((1 - \gamma) \cdot p_t^{dx} + \gamma/k\right) \cdot p^x,$$

where E_t denotes the conditional expectation given observations up to wave $t - 1$, and conditional on $\boldsymbol{\theta}$. We can rewrite the sample average as

$$\bar{Y}_t^{dx} = \frac{\sum_{i,t' \leq t} W_{it'} Y_{it'}}{\sum_{i,t' \leq t} Z_{it'}} \cdot \frac{\sum_{i,t' \leq t} Z_{it'}}{\sum_{i,t' \leq t} W_{it'}}.$$

We have by construction that $Z_{it} \geq p^x \cdot \gamma/k$, and since $N_t \geq \underline{N}$, it follows that $\sum_{i,t' \leq t} Z_{it'} \to \infty$ as $t \to \infty$. Applying Lemma C.C.1 to the first fraction, and a standard law of large numbers to the inverse of the second fraction, we get that

$$\bar{Y}_t^{dx} \to \theta_0^{dx}$$

in probability as $t \to \infty$.

1. Given the assumed uniqueness of d^{*x}, there exists an ε-neighborhood of $\boldsymbol{\theta}_0$ such that d^{*x} is constant for all x in this neighborhood. The claim follows if we can show that the posterior probability of such an ε-neighborhood goes to 1 in probability as $t \to \infty$.

    Given our assumption that the prior for $\boldsymbol{\theta}$ has full support, this condition follows from Schwartz's theorem (Theorem C.1), if we can show existence of a consistent test for the hypothesis that $\boldsymbol{\theta} = \boldsymbol{\theta}_0$ against the alternative that $\|\boldsymbol{\theta} - \boldsymbol{\theta}_0\| > \varepsilon$.

    In our setting such a test can be constructed by setting

    $$\varphi_t = \mathbf{1}\left(\|\bar{Y} - \boldsymbol{\theta}_0\| > \varepsilon/2\right).$$

    The required consistency follows by convergence in probability of $\bar{Y}$.

2. By construction of our algorithm, treatment d is assigned with probability $(1 - \gamma) \cdot p_t^{dx} + \gamma/k$ to units in stratum x in period t. It follows from item 1 that this probability converges to $\bar{q}^{dx}$ as $t \to \infty$.

    Since N_t is bounded below, the same holds for the cumulative share $\bar{q}_t^{dx}$.

3. By definition,

$$Regret_t = \sum_{x,d} \Delta^{dx} \bar{q}_t^{dx} \bar{p}_t^x,$$

where $\bar{p}_t^x$ is the share of observations in stratum x up to period t. The claim follows from item 2, and the law of large numbers for $\bar{p}_t^x$, once we note that $\Delta^{dx} = 0$ for $d = d^{*x}$.

4. This is an immediate consequence of Corollary 2.1 and Theorem 3.2 in Melfi and Page (2000), as re-stated in our Theorem C.2 above. The necessary condition of their Corollary 2.1 follows immediately from the fact that our Tempered Thompson Algorithm is adapted to the filtration induced by the history of data. The necessary conditions of their Theorem 3.2 (converging shares) are verified by our item 2 above. □

Appendix D: Markov Chain Monte Carlo

Algorithm 1 Markov Chain Monte Carlo for the hierarchical Bayes model

---

Require: The cumulated assignment frequencies m^{dx} and success numbers r^{dx}.
  Starting values $\boldsymbol{\alpha}_0, \boldsymbol{\beta}_0$, length of the burn in period B, and number of draws R.
1: **for** $\rho = 1$ **to** $B + R$ **do**
2:    Gibbs step:
      Given $\boldsymbol{\alpha}_{\rho-1}$ and $\boldsymbol{\beta}_{\rho-1}$, for all d, x
      draw θ^{dx} from the $Beta(\alpha_\rho^d + r^{dx}, \beta_\rho^d + m^{dx} - r^{dx})$ distribution.
3:    Metropolis step 1:
      Given $\boldsymbol{\beta}_{\rho-1}$ and $\boldsymbol{\theta}_\rho$, draw α_ρ^d
      by sampling from a normal proposal distribution (truncated below).
      Accept this draw if an independent uniform draw is less than the ratio of the posterior for the new draw, relative to the posterior for $\alpha_{\rho-1}^d$.
      Otherwise set $\alpha_\rho^d = \alpha_{\rho-1}^d$.
4:    Metropolis step 2:
      Similarly for $\boldsymbol{\beta}_{\rho-1}$ given $\boldsymbol{\theta}_\rho$ and $\boldsymbol{\alpha}_{\rho-1}$.
5: **end for**
6: Throw away all draws from the burn-in period $\rho = 1, \ldots, B$.
7: **return** For all x and d, the estimated probabilities

$$\hat{p}^{dx} = \frac{1}{R} \sum_{\rho=B+1}^{B+R} \mathbf{1}\left(d = \arg\max_{d'} \theta_\rho^{d'x}\right). \tag{D.1}$$

Denote by $\boldsymbol{\theta}, \boldsymbol{m}_t, \boldsymbol{r}_t$ the vectors of parameters, cumulative trials, and cumulative successes, where each of these is indexed by both d and x, and denote by $\boldsymbol{\alpha}, \boldsymbol{\beta}$ the vectors of hyperparameters indexed by d. Let ρ index replication draws, with ρ ranging from 1 to $B + R$. We sample from the posterior distribution of $(\boldsymbol{\theta}, \boldsymbol{\alpha}, \boldsymbol{\beta})$ given $\boldsymbol{m}_{t-1}, \boldsymbol{r}_{t-1}$ using the Markov Chain Monte Carlo algorithm described in Algorithm 1. Markov Chain Monte Carlo methods are reviewed in Gelman et al. (2014), chapter 11.

Algorithm 1 converges to a stationary distribution that equals the joint posterior of $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ given $\boldsymbol{m}_t, \boldsymbol{r}_t$. In particular, we have that the posterior probability that a treatment d is optimal given x, in the sense that it maximizes the probability of employment, is given by

$$p_t^{dx} = P\left(d = \arg\max_{d'} \theta^{d'x} | \boldsymbol{m}_t, \boldsymbol{r}_t\right) = \plim_{R\to\infty} \frac{1}{R} \sum_{\rho=1}^{R} \mathbf{1}\left(d = \arg\max_{d'} \theta_\rho^{d'x}\right).$$
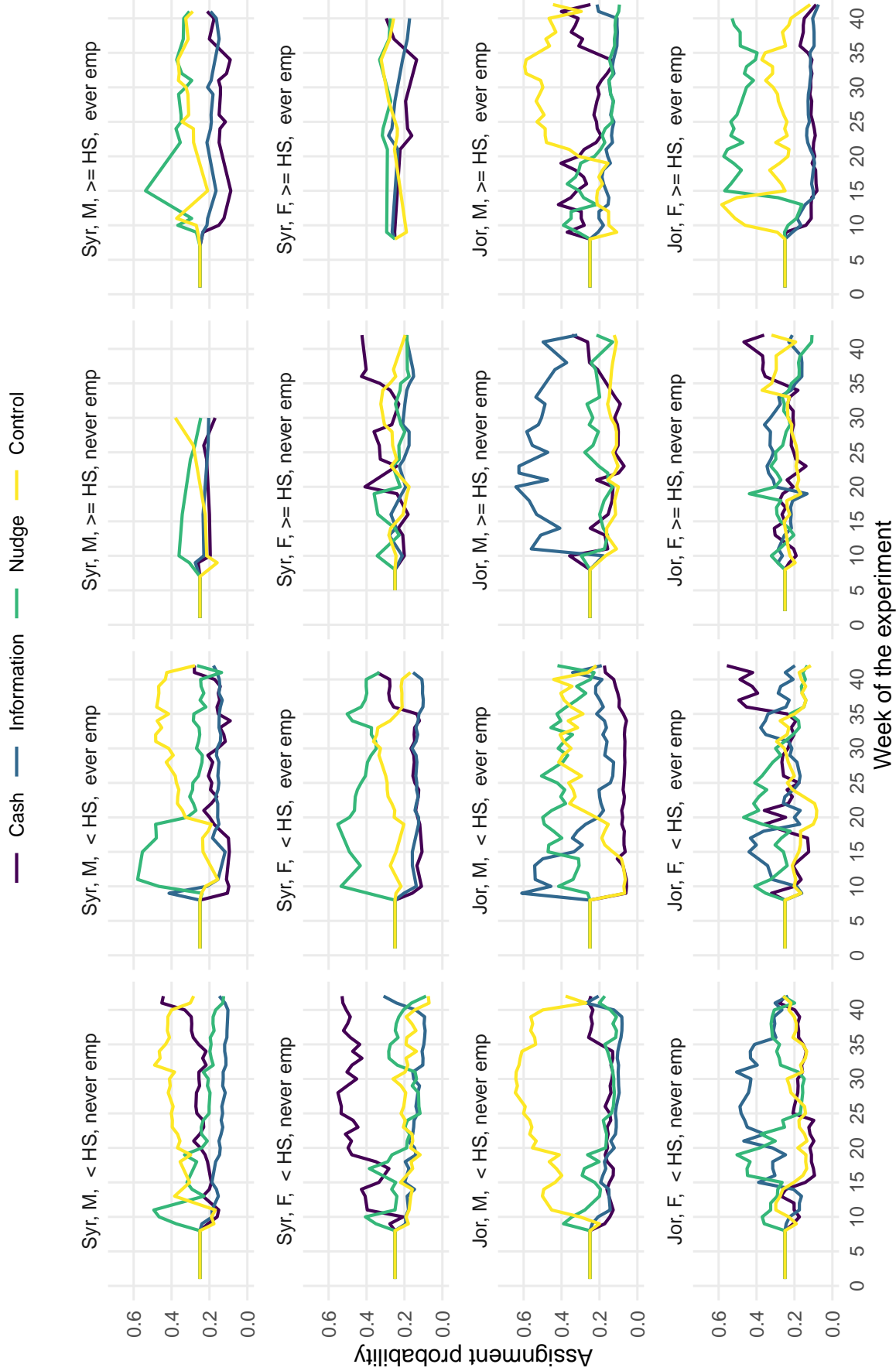(D.2)

In our implementation of this algorithm, we use a warm-up period of $B = 1,000$, and then draw $R = 10,000$ replications; averaging over these gives our estimated posterior distribution. These values are generously chosen relative to standard recommendations (cf. Gelman et al. (2014) chapter 11), making convergence likely. In our simulations these values yield stable posterior probabilities.

Appendix E: Additional Table and Figures

TABLE E.A.1. **95% credible sets for average potential outcomes**

stratum	Cash	Information	Nudge	Control
Syr, M, < HS, never emp	(0.010, 0.110)	(0.000, 0.080)	(0.010, 0.090)	(0.010, 0.100)
Syr, M, < HS, ever emp	(0.030, 0.120)	(0.010, 0.090)	(0.030, 0.100)	(0.030, 0.100)
Syr, M, >= HS, never emp	(0.020, 0.260)	(0.000, 0.170)	(0.010, 0.140)	(0.020, 0.240)
Syr, M, >= HS, ever emp	(0.010, 0.170)	(0.000, 0.170)	(0.020, 0.150)	(0.010, 0.180)
Syr, F, < HS, never emp	(0.010, 0.050)	(0.010, 0.060)	(0.000, 0.050)	(0.000, 0.030)
Syr, F, < HS, ever emp	(0.010, 0.080)	(0.010, 0.110)	(0.020, 0.080)	(0.010, 0.070)
Syr, F, >= HS, never emp	(0.020, 0.190)	(0.000, 0.150)	(0.020, 0.150)	(0.000, 0.140)
Syr, F, >= HS, ever emp	(0.010, 0.180)	(0.000, 0.160)	(0.010, 0.130)	(0.000, 0.160)
Jor, M, < HS, never emp	(0.010, 0.110)	(0.010, 0.090)	(0.020, 0.120)	(0.030, 0.120)
Jor, M, < HS, ever emp	(0.030, 0.170)	(0.040, 0.150)	(0.050, 0.140)	(0.060, 0.160)
Jor, M, >= HS, never emp	(0.040, 0.230)	(0.040, 0.220)	(0.020, 0.150)	(0.000, 0.140)
Jor, M, >= HS, ever emp	(0.030, 0.150)	(0.020, 0.160)	(0.010, 0.110)	(0.040, 0.150)
Jor, F, < HS, never emp	(0.010, 0.070)	(0.020, 0.080)	(0.030, 0.090)	(0.010, 0.080)
Jor, F, < HS, ever emp	(0.060, 0.190)	(0.050, 0.170)	(0.020, 0.100)	(0.030, 0.130)
Jor, F, >= HS, never emp	(0.030, 0.150)	(0.010, 0.100)	(0.010, 0.080)	(0.020, 0.130)
Jor, F, >= HS, ever emp	(0.000, 0.110)	(0.000, 0.100)	(0.060, 0.180)	(0.020, 0.110)

Note: The table reports results for wage employment at the time of the rapid follow-up interview.

FIGURE E.A.1. **Assignment probabilities by stratum and by week**

References

Abebe, Girum, A Stefano Caria, and Esteban Ortiz-Ospina (2021). "The Selection of Talent: Experimental and Structural Evidence from Ethiopia." *American Economic Review*, 111(6), 1757–1806.

Abebe, Girum Tefera, Stefano Caria, Marcel Fafchamps, Paolo Falco, Simon Franklin, and Simon Quinn (2020). "Anonymity or Distance? Job Search and Labour Market Exclusion in a Growing African City."

Abel, Martin, Rulof Burger, Eliana Carranza, and Patrizio Piraino (2019). "Bridging the Intention-behavior Gap? The Effect of Plan-making Prompts on Job Search and Employment." *American Economic Journal: Applied Economics*, 11(2), 284–301.

Abu Hamad, Bassam, Nicola Jones, Fiona Samuels, Ingrid Gercama, Elizabeth Presler-Marshall, Georgia Plank, Aida Essaid, Said Ebbini, Kifah Bani Odeh, Deya'eddin Bazadough, Hala Abu Taleb, Hadeel Al Amayreh, and Jude Sadji (2017). "A Promise of Tomorrow: The effects of UNHCR and UNICEF cash assistance on Syrian refugees in Jordan." *Technical report: https://www.unhcr.org/5a1d1b7b4.pdf* .

Agrawal, Shipra and Navin Goyal (2012). "Analysis of thompson sampling for the multi-armed bandit problem." In *Conference on Learning theory*, pp. 39–1.

Agrawal, Shipra and Navin Goyal (2013). "Further Optimal Regret Bounds for Thompson Sampling." In *Artificial intelligence and statistics*, pp. 99–107.

Alfonsi, Livia, Oriana Bandiera, Vittorio Bassi, Robin Burgess, Imran Rasul, Munshi Sulaiman, and Anana Vitali (2020). "Tackling Youth Unemployment: Evidence from a Labor Market Experiment in Uganda." *Econometrica*.

Alrababa'h, Ala', Scott Williamson, Andrea Balacar Dillon, Dominik Hangartner, and Jens Hainmueller (2019). "Attitudes toward migrants in a highly impacted economy: evidence from the Syrian refugee crisis in Jordan." *Comparative Political Studies*.

Andrews, Isaiah, Toru Kitagawa, and Adam McCloskey (2019). "Inference on winners." Tech. rep., cemmap working paper.

Athey, Susan, Raj Chetty, Guido W Imbens, and Hyunseung Kang (2019). "The Surrogate Index: Combining Short-term Proxies to Estimate Long-term Treatment Effects more Rapidly and Precisely." Working Paper 26463, National Bureau of Economic Research.

Athey, Susan and Guido W Imbens (2017). "The Econometrics of Randomized Experiments." In *Handbook of Economic Field Experiments*, vol. 1, pp. 73–140. Elsevier.

Attanasio, Orazio, Adriana Kugler, and Costas Meghir (2011). "Subsidizing vocational training for disadvantaged youth in Colombia: Evidence from a randomized trial." *American Economic Journal: Applied Economics*, 3(3), 188–220.

Balboni, Clare, Oriana Bandiera, Robin Burgess, Maitreesh Ghatak, and Anton Heil (2020). "Why do People Stay Poor?" *CEPR Discussion Paper No. DP14534*.

Banerjee, Abhijit, Rukmini Banerji, James Berry, Esther Duflo, Harini Kannan, Shobhini Mukerji, Marc Shotland, and Michael Walton (2017a). "From Proof of Concept to Scalable Policies: Challenges and Solutions, with an Application." *Journal of Economic Perspectives*, 31(4), 73–102.

Banerjee, Abhijit V, Rema Hanna, Gabriel E Kreindler, and Benjamin A Olken (2017b). "Debunking the Stereotype of the Lazy Welfare Recipient: Evidence from Cash Transfer Programs." *The World Bank Research Observer*, 32(2), 155–184.

Banerjee, Abhijit V and Andrew F Newman (1993). "Occupational Choice and the Process of Development." *Journal of Political Economy*, 101(2), 274–298.

Banerjee, Abhijit V and Sandra Sequeira (2020). "Spatial Mismatches and Imperfect Information in the Job Search." *CEPR Discussion Paper No. DP14414*.

Barbelet, Veronique, Jessica Hagen-Zanker, and Dina Mansour-Ille (2018). "The Jordan Compact: Lessons Learnt and Implications for Future Refugee Compacts." https://www.odi.org/sites/odi.org.uk/files/resource-documents/12058.pdf.

Bassi, Vittorio and Aisha Nansamba (2020). "Information Frictions in the Labor Market: Evidence from a Field Experiment in Uganda." *Working Paper*.

Battisti, Michele, Yvonne Giesing, and Nadzeya Laurentsyeva (2019). "Can Job Search Assistance Improve the Labour Market Integration of Refugees? Evidence from a Field Experiment." *Labour Economics*, 61, 101745.

Benhassine, Najy, Florencia Devoto, Esther Duflo, Pascaline Dupas, and Victor Pouliquen (2015). "Turning a Shove into a Nudge? A 'Labeled Cash Transfer' for Education." *American Economic Journal: Economic Policy*, 7(3), 86–125.

Berry, Donald A (2006). "Bayesian clinical trials." *Nature Reviews: Drug discovery*, 5(1), 27–36.

Betts, Alexander and Paul Collier (2017). *Refuge: Transforming a Broken Refugee System*. Penguin UK.

Bhattacharya, Debopam and Pascaline Dupas (2012). "Inferring welfare maximizing treatment assignment under budget constraints." *Journal of Econometrics*, 167(1), 168–196.

Bryan, Gharad, Shyamal Chowdhury, and Ahmed Mushfiq Mobarak (2014). "Underinvestment in a profitable technology: The case of seasonal migration in Bangladesh." *Econometrica*, 82(5), 1671–1748.

Bubeck, Sébastien and Nicolò Cesa-Bianchi (2012). "Regret Analysis of Stochastic and Nonstochastic Multi-armed Bandit Problems." *Foundations and Trends® in Machine Learning*, 5(1), 1–122.

Caliendo, Marco, Deborah A Cobb-Clark, and Arne Uhlendorff (2015). "Locus of control and job search strategies." *Review of Economics and Statistics*, 97(1), 88–103.

Cesa-Bianchi, Nicolò, Roberto Colomboni, and Maximilian Kasy (2022). "Adaptive maximization of social welfare." *Working Paper*.

Cesa-Bianchi, Nicolo and Gábor Lugosi (2006). *Prediction, learning, and games*. Cambridge university press.

Chaaban, Jad, Hala Ghattas, Nisreen Salti, Wael Moussa, Alexandra Irani, Zeina Jamaluddine, and Rima Al Mokdad (2020). "Impact Evaluation of the Well-being of

Syrian Refugees." *Technical report: https://reliefweb.int/sites/reliefweb.int/files/resources/77377.pdf*.

Chamberlain, Gary (2011). "Bayesian Aspects of Treatment Choice." *The Oxford Handbook of Bayesian Econometrics*, pp. 11–39.

Chernozhukov, Victor, Mert Demirer, Esther Duflo, and Ivan Fernandez-Val (2018). "Generic machine learning inference on heterogenous treatment effects in randomized experiments." Tech. rep., National Bureau of Economic Research.

Deaton, Angus (2020). "Randomization in the Tropics Revisited: A Theme and Eleven Variations." Tech. Rep. 27600, National Bureau of Economic Research.

Dehejia, Rajeev H (2005). "Program Evaluation as a Decision Problem." *Journal of Econometrics*, 125(1-2), 141–173.

DellaVigna, Stefano and M Daniele Paserman (2005). "Job Search and Impatience." *Journal of Labor Economics*, 23(3), 527–588.

Devictor, Xavier and Quy-Toan Do (2017). "How Many Years Have Refugees Been in Exile?" *Population and Development Review*, 43(2), 355–369.

Duflo, Esther and Abhijit Banerjee (2017). *Handbook of Field Experiments*. Elsevier.

Fasani, Francesco, Tommaso Frattini, and Luigi Minale (2021). "Lift the ban? Initial employment restrictions and refugee labour market outcomes." *Journal of the European Economic Association*, 19(5), 2803–2854.

FDA, Center for Drug Evaluation, Center for Biologics Evaluation Research (CDER), and Research (CBER) (2018). "Adaptive Designs for Clinical Trials of Drugs and Biologics, Guidance for Industry." *U.S. Department of Health and Human Services, Food and Drug Administration*.

Franklin, Simon (2018). "Location, Search Costs and Youth Unemployment: Experimental Evidence from Transport Subsidies." *The Economic Journal*, 128(614), 2353–2379.

Gelman, Andrew, John B Carlin, Hal S Stern, and Donald B Rubin (2014). *Bayesian data analysis*, vol. 2. Taylor & Francis.

Ghosal, Subhashis and Aad Van der Vaart (2017). *Fundamentals of Nonparametric Bayesian Inference*, vol. 44. Cambridge University Press.

Gittins, John C (1979). "Bandit Processes and Dynamic Allocation Indices." *Journal of the Royal Statistical Society: Series B (Methodological)*, 41(2), 148–164.

Gordon, Grant (2017). "Solving the Refugee Employment Problem in Jordan: A Survey of Syrian Refugees." *International Rescue Committee report*.

Government of Jordan (2019). "Jordan Response Plan for the Syrian Crisis: Draft manuscript."

Groh, Matthew, Nandini Krishnan, David McKenzie, and Tara Vishwanath (2016a). "Do Wage Subsidies Provide a Stepping-stone to Employment for Recent College Graduates? Evidence from a Randomized Experiment in Jordan." *Review of Economics and Statistics*, 98(3), 488–502.

Groh, Matthew, Nandini Krishnan, David McKenzie, and Tara Vishwanath (2016b). "The Impact of Soft Skills Training on Female Youth Employment: Evidence from a Randomized Experiment in Jordan." *IZA Journal of Labor & Development*, 5(1), 9.

Groh, Matthew, David McKenzie, Nour Shammout, and Tara Vishwanath (2015a). "Testing the Importance of Search Frictions and Matching Through a Randomized Experiment in Jordan." *IZA Journal of Labor Economics*, 4(1), 1–20.

Groh, Matthew, David McKenzie, and Tara Vishwanath (2015b). "Reducing Information Asymmetries in the Youth Labor Market of Jordan with Psychometrics and Skill Based Tests." *The World Bank Economic Review*, 29(suppl_1), S106–S117.

Hadad, Vitor, David A. Hirshberg, Ruohan Zhan, Stefan Wager, and Susan Athey (2019). "Confidence Intervals for Policy Evaluation in Adaptive Experiments." *Working paper*.

Hagen-Zanker, Jessica, Martina Ulrichs, and Rebecca Holmes (2018). "What are the Effects of Cash Transfers for Refugees in the Context of Protracted Displacement? Findings from Jordan." *International Social Security Review*, 71(2), 57–77.

Haushofer, Johannes and C Jessica E Metcalf (2020). "Which Interventions Work Best in a Pandemic?" *Science*, 368(6495), 1063–1065.

Haushofer, Johannes and Jeremy Shapiro (2016). "The Short-Term Impact of Unconditional Cash Transfers to the Poor: Experimental Evidence from Kenya." *The Quarterly Journal of Economics*, 131(4), 1973–2042.

Herkenhoff, Kyle, Gordon Phillips, and Ethan Cohen-Cole (2016). "How Credit Constraints Impact Job Finding Rates, Sorting & Aggregate Output." Tech. rep., National Bureau of Economic Research.

Hussam, Reshmaan N, Erin M Kelley, Gregory V Lane, and Fatima T Zahra (2021). "The Psychosocial Value of Employment." *NBER Working Paper No. 28924*.

IRC (2017). "In Search of Work: Creating Jobs for Syrian Refugees: A Case Study of the Jordan Compact." https://www.rescue.org/sites/default/files/document/1343/insearchofworkweb.pdf.

Kasy, Maximilian (2015). "Partial Identification, Distributional Preferences, and the Welfare Ranking of Policies." *Review of Economics and Statistics*.

Kasy, Maximilian and Anja Sautmann (2021). "Adaptive treatment assignment in experiments for policy choice." *Econometrica*, 89(1), 113–132.

Kasy, Maximilian and Alexander Teytelboym (2020). "Adaptive Targeted Infectious Disease Testing." *Oxford Review of Economic Policy*.

Kasy, Maximilian and Alexander Teytelboym (2023). "Matching with semi-bandits." *Econometrics Journal*.

Kaufmann, Emilie, Nathaniel Korda, and Rémi Munos (2012). "Thompson Sampling: An Asymptotically Optimal Finite-time Analysis." In *International Conference on Algorithmic Learning Theory*, pp. 199–213, Springer.

Kitagawa, Toru and Aleksey Tetenov (2018). "Who Should be Treated? Empirical Welfare Maximization Methods for Treatment Choice." *Econometrica*, 86(2), 591–616.

Kluve, Jochen, Susana Puerto, David Robalino, Jose M Romero, Friederike Rother, Jonathan Stöterau, Felix Weidenkaff, and Marc Witte (2019). "Do Youth Employment Programs Improve Labor Market Outcomes? A Quantitative Review." *World Development*, 114, 237–253.

Lattimore, Tor and Csaba Szepesvári (2020). *Bandit algorithms*. Cambridge University Press.

Lehmann, Christian and Daniel Masterson (2014). "Emergency Economies: The Impact of Cash Assistance in Lebanon." *Technical report:* *https://www.rescue.org/sites/default/files/document/631/ emergencyeconomiesevaluationreport-lebanon2014.pdf*.

Lewi, Yuval, Haim Kaplan, and Yishay Mansour (2020). "Thompson Sampling for Adversarial Bit Prediction." In *Algorithmic Learning Theory*, pp. 518–553, PMLR.

Loiacono, Francesco and Mariajose Silva Vargas (2019). "Improving Access To Labour Markets for Refugees: Evidence from Uganda." *Working paper*.

McKenzie, David (2017). "How Effective are Active Labor Market Policies in Developing Countries? A Critical Review of Recent Evidence." *The World Bank Research Observer*, 32(2), 127–154.

Meager, Rachael (2019). "Understanding the Average Impact of Microcredit Expansions: A Bayesian Hierarchical Analysis of Seven Randomized Experiments." *American Economic Journal: Applied Economics*, 11(1), 57–91.

Meghir, Costas, Ahmed Mushfiq Mobarak, Corina D Mommaerts, and Melanie Morten (2019). "Migration and Informal Insurance: Evidence from a Randomized Controlled Trial and a Structural Model." Tech. rep., National Bureau of Economic Research.

Melfi, Vincent F and Connie Page (2000). "Estimation After Adaptive Allocation." *Journal of Statistical Planning and Inference*, 87(2), 353–363.

Muralidharan, Karthik and Paul Niehaus (2017). "Experimentation at Scale." *Journal of Economic Perspectives*, 31(4), 103–24.

REACH (2021). "Likelihood Assessment of Micro Businesses: Assessment of Home-Based Businesses and Other Self-Employment." *Report*.

Russell, Stuart (2019). *Human Compatible: Artificial Intelligence and the Problem of Control*. Penguin.

Russo, Daniel (2016). "Simple Bayesian Algorithms for Best Arm Identification." In *Conference on Learning Theory*, pp. 1417–1418.

Russo, Daniel and Benjamin Van Roy (2016). "An information-theoretic analysis of thompson sampling." *The Journal of Machine Learning Research*, 17(1), 2442–2471.

Russo, Daniel J, Benjamin Van Roy, Abbas Kazerouni, Ian Osband, Zheng Wen, et al. (2018). "A Tutorial on Thompson Sampling." *Foundations and Trends in Machine Learning*, 11(1), 1–96.

Schuettler, Kirsten and Laura Caron (2020). "Jobs Interventions for Refugees and Internally Displaced Persons." *World Bank Jobs Working Paper No. 47*.

Shami, Soha (2019). "When Worlds Collide: Lessons learned from the intersection of behavioral and human-centered design in humanitarian contexts." *https://medium.com/airbel/ lessons-learned-from-the-intersection-of-behavioral-and-human-centered-design-in-humanitarian*

Slivkins, Aleksandrs (2019). "Introduction to multi-armed bandits." *arXiv preprint arXiv:1904.07272*.

Taylor, Jessica, Eliezer Yudkowsky, Patrick LaVictoire, and Andrew Critch (2016). "Alignment for Advanced Machine Learning Systems." Tech. rep., Machine Intelligence Research Institute.

Thompson, William R. (1933). "On the likelihood that one unknown probability exceeds another in view of the evidence of two samples." *Biometrika*, 25(3/4), 285–294.

Tiltnes, Åge A, Huafeng Zhang, and Jon Pedersen (2019). "The living conditions of Syrian refugees in Jordan." *FAFO Report. Amman: Fafo Research Foundation*.

UNHCR (2017). "Setting up a Home Based business & other Businesses; basic facts from AMMAN and Al-Mafraq." *Presentation, UNHCR Legal Unit*.

UNHCR (2020a). "Syria Emergency." https://www.unhcr.org/syria-emergency.html.

UNHCR (2020b). "UNHCR: Figures at a Glance." https://www.unhcr.org/en-us/figures-at-a-glance.html.

van der Vaart, Aad W. (2000). *Asymptotic statistics*. Cambridge University Press.

Verme, Paolo, Chiara Gigliarano, Christina Wieser, Kerren Hedlund, Marc Petzoldt, and Marco Santacroce (2015). *The Welfare of Syrian Refugees: Evidence from Jordan and Lebanon*. The World Bank.

Wager, Stefan and Susan Athey (2018). "Estimation and Inference of Heterogeneous Treatment Effects using Random Forests." *Journal of the American Statistical Association*, 113(523), 1228–1242.

Wager, Stefan and Kuang Xu (2021). "Diffusion Asymptotics for Sequential Experiments." *arXiv preprint arXiv:2101.09855*.

Wang, Siwei and Wei Chen (2018). "Thompson Sampling for Combinatorial Semi-Bandits." In *Proceedings of the 35th International Conference on Machine Learning*, *Proceedings of Machine Learning Research*, vol. 80, edited by Jennifer Dy and Andreas Krause, pp. 5114–5122, PMLR, URL https://proceedings.mlr.press/v80/wang18a.html.

Weber, Richard et al. (1992). "On the Gittins index for multiarmed bandits." *The Annals of Applied Probability*, 2(4), 1024–1033.