Only You: A Field Experiment of Text Message to Prevent Free-riding in Japan Marrow Donor Program

Hiroki Kato (Hitostubashi Institute for Advanced Study, Hitotsubashi University, Tokyo, Japan)

Fumio Ohtake (Center for Infectious Disease Education and Research (CiDER), Osaka University, Osaka, Japan)

Saiko Kurosawa (Department of Oncology, Ina Central Hospital, Nagano, Japan)

Kazuhiro Yoshiuchi (Graduate School of Medicine, Tokyo University, Tokyo, Japan)

Takahiro Fukuda (Department of Hematopoietic Stem Cell Transplantation, National Cancer Center Hospital, Tokyo, Japan)

Last updated on June 20, 2023

Abstract

日本骨髄バンクを介した同種造血幹細胞移植は移植に至るまで長時間を要し、登録患者の約半数しか移植を受けられない。これは移植コーディネーションの多くは登録ドナーが提供に消極的であることによって移植実施前に中断してしまうからである。本研究は日本骨髄バンクの協力のもと、登録ドナーの提供意向を高めるような情報を提供する介入をフィールド実験にて検証する。我々は11,049名の適合ドナーのコーディネーションデータを用いて、情報提供の効果を推定した。我々は患者一人あたりにマッチする潜在ドナーの数が少ないという情報が20代男性の提供意向を25%高めることを明らかにした。また、早いコーディネーションが患者の移植率を高めるという情報が20代女性の早期返信を促していることも明らかにした。これらの結果は情報提供が特定の性・年代にのみ影響を与えており、特に、移植成績の良い若年ドナーの行動変容を促していることを示唆している。

# 1 Introduction

同種造血幹細胞移植は白血病などの血液病に対する最も再発率の低い治療法の一つである。この治療法では、(1)抗がん剤や放射線によって異常な細胞と健康な造血幹細胞を同時に殺し、(2)他者から提供された健康な造血幹細胞を移植する。骨髄移植（造血幹細胞移植）は、ドナーのHLAと呼ばれる白血球の型が患者のHLAと一致していることが条件である[[1]](#footnote-20)。ランダムに選ばれた二人のマッチング確率は1%未満である一方で、兄弟姉妹間で一致する可能性が最も高く、その確率は30%程度である。また、親子間のマッチング確率はかなり低い。親族の中に適合するドナーがいなければ、患者は近親者以外からドナーを探さなければならない。日本において、患者は一般的に日本骨髄バンク（JMDP）を介して非近親者のドナーを探すことになる。しかしながら、JMDPによるコーディネーションは移植に至るまで長時間を要し、登録患者の60%のみが移植を受けている (Hirakawa et al. 2018)。したがって、移植までの期間を短くし、登録患者の移植到達率を高めることが重要である。

移植率を高めるためのドナープールへの介入策は二種類ある。第一に、潜在ドナーの人数を増やして、マッチング確率を高めるような政策である。4人未満のドナーと合致する患者と比較して、200人以上のドナーと合致する患者の移植率は45%から74%に上昇する(Hirakawa et al. 2018)。しかしながら、Takanashi (2016) によると、2000年から2015年にかけて潜在的なドナーの数は約二倍に増えたにもかかわらず、初回マッチング確率は5%程度しか増えていない。潜在ドナーを増やすことの限界便益は小さい[[2]](#footnote-21)。

第二の政策は提供に積極的な潜在ドナーの比率を高め、ドナープールの質を高める。Hirakawa et al. (2018) はコーディネーションの多く（2004–2013年で実施されたコーディネーションの73%）はコーディネーションの第一過程である確認検査の前に中断しており、その多くは（不健康を含む）ドナー側の理由であることを明らかにしている。特に、若年ドナーほど健康理由でコーディネーションを中断する可能性が低く、それ以外の自身の理由でコーディネーションを中断する可能性が高い[[3]](#footnote-22)。これはJMDPだけでなく他国の骨髄ドナープログラムでも問題となっている(Haylock et al. 2022)。不健康以外のドナー理由が本人の意向と相関しているのであれば、強い提供意向がある潜在ドナーを増やすことは移植到達率を高め、その限界便益は潜在ドナーを増やすことの限界便益よりも大きいだろう。

そこで、本研究はドナープールの質を高める施策の一つとして、提供意向を高めるような情報提供の効果を検証する。JMDPに登録した人がある患者とマッチングすると、適合ドナーはJMDPから適合通知を受け取る。そして、提供意思を示して適合通知に返信する潜在ドナーは移植のためのコーディネーションを受ける。我々はJMDPによって公開されている情報に基づいて適合通知に二つの新規のメッセージを加え、追加メッセージの効果を検証するフィールド実験をJMDPと協働で実施した。

第一の新規メッセージは患者一人あたりのHLA適合ドナーの数が少ないという情報を提示した。同じHLA型を持つ他の潜在ドナーがプールに存在すれば、自身の提供はその適合ドナーと代替可能であり、自身の提供なしに患者の生存による効用を得られる。また、一人の患者に対して、複数のドナー（最大10人）がコーディネーションを同時に受ける。したがって、JMDPを介した移植は公共財の性質を有しており、一般的なHLAの型を持っている人ほど提供に消極的となるだろう。すなわち、JMDPを介した移植は標準的なフリーライド問題に直面している(Bergstrom, Garratt, and Sheehan-Connor 2009)。

この問題は1人だけの協力行動で公共財が生産される状況を想定するボランティアのジレンマと本質的に同じである。ボランティアのジレンマでは、1人でも協力行動を取る確率がグループサイズによって減少することが理論的に予測され、その仮説は実験室実験でも確認されている(Diekmann 1985, 1986; Franzen 1999; Davis and Heller 2017)。また、JMDPのコーディネーションを経験した人に対するインタビュー研究(Kurosawa et al. 2022)によると、提供意向の低い人は「何人かいるドナーの1人」という意識を持っており、自身の提供が他者によって代替可能であることが提供意向の阻害要因であることを示唆している。ボランティアのジレンマとの重要な違いはJMDPの潜在ドナーは自身のHLA型や同時にコーディネーションを受けている人の数について知ることができないので、グループサイズを正確に把握できないことにある。したがって、潜在ドナーは代替可能な他のドナーの数を過度に多く見積もり、提供に消極的になる可能性がある。一つ目のメッセージはこの誤解から生じる行動を修正することを目的としている。

第二のメッセージは早いコーディネーションが患者の移植到達率を高めるという情報を提示した。この情報は患者への移植によって得られる効用が経時的に減衰していくことを暗に提示している。もしその効用水準が時間に対して一定であると誤解しているならば、今コストを支払ってコーディネーションを受けるよりも将来コストを支払ってコーディネーションを受けようと考えるはずである。同じ意思決定が将来でも生じ、適合通知への返信が遅れてしまう可能性がある。こうした問題は行動経済学における現在バイアスによる先延し行動と似たような問題である(O’Donoghue and Rabin 2001)。先延ばし行動は将来の自分が価値を合理的に評価できると信じているが、実際はその時の効用を過剰に高く評価してしまうことで生じる。二つ目のメッセージは患者への移植によって得られる効用水準の経時的変化に関する誤解から生じる先延ばし行動を修正することを目的としている。

我々は以上のメッセージの効果を検証するために、四つの実験群を作成した。一つ目の実験群は上記の二つのメッセージを追加していない標準的な適合通知を送付する（コントロール群）。第二・三の実験群はそれぞれ上記のメッセージを一つ追加した適合通知を送付する。第四の実験群は二つのメッセージを同時に追記した適合通知を送付する。我々は2021年9月から2022年2月の間に適合通知を受け取った11,154名の適合ドナーを対象にフィールド実験を実施し、実験群を週単位のクラスターランダム化で実験参加者に割り当てた。我々はJMDPの協力のもと、コーディネーションデータを2022年6月末に受け取り、メッセージの効果を検証した。

実験の結果、メッセージの効果は特定の性・年代にのみ影響を与えていることを発見した。本研究の主要な結果は二つある。第一に、患者一人当たりのHLA適合ドナーが少ないという情報のみを提供することは20代男性の提供意向を高めたが、他の性・年代では提供意向に統計的に有意な影響を与えていない。第二に、ドナーが早く見つかるほど移植到達率を高められるという情報を与えることは20代女性の全体の返信率に影響を与えていないが、短期間の返信（4日以内）に正の影響を与えている[[4]](#footnote-23)。すなわち、この情報は返信行動そのものを促進しているよりむしろ、返信日数を短くしているといえる。このような効果は他の性・年代では確認されなかった。

本研究は情報提供によって骨髄バンクに登録した潜在ドナーの意向を高められることを示唆するものであり、これはJMDPを含めた世界中の骨髄ドナープログラムの実務的な洞察をもたらすものである。JMDPと同様に、ドイツを中心とする国際的な骨髄ドナープログラムであるDKMSやアメリカの骨髄ドナープログラムであるNMDPは順調に登録者を増やしている一方で、登録者のモチベーションが維持されずコーディネーションに至らないことが課題となっている(G. E. Switzer et al. 1999; Galen E. Switzer et al. 2004; Haylock et al. 2022)。これまでの研究は、ドナー休暇の法制化(Lacetera, Macis, and Stith 2014)・ドナーのモチベーション維持に資するDKMS独自の施策(Haylock et al. 2022)などの効果を検証してきた。Galen E. Switzer et al. (2018) はNMDPがマッチした潜在ドナーに提供を依頼するときに、NMDPが送るメッセージに介入を施した。彼らの介入は「あなたは患者のHLA型と完全に一致するという特別な状況にいる」というメッセージであり、これをHLAが完全一致した潜在ドナーに電話で伝えた。彼らの実験は完全なランダム比較試験ではなく、この新規メッセージがコーディネーションの数を増やさないことを明らかにした。我々の介入はこの研究と非常に近いものである。本研究では、メッセージ効果はかなり異質的であり、特定の性・年代にのみ影響を与えていることを示す。

また、本研究は献血などの経済学研究に貢献するものである。造血幹細胞移植と献血は共通点があるものの、いくつかの違いがある。献血は提供の意向を示す時点と実際の提供の時点が同じであるが、造血幹細胞移植は二つの時点が異なっている。そのため、先に述べたように、潜在ドナーの意向を維持することが骨髄ドナープログラムの課題となっている。献血の文脈では、金銭的インセンティブやSymbolic rewardなどの非金銭的インセンティブ効果検証を通じて、献血の動機が検証されている(Mellström and Johannesson 2008; Lacetera and Macis 2010)。また、造血幹細胞移植と同じように、献血は公共財とみなせる。Wildman and Hollingsworth (2009) は献血者のみのデータを用いて、フリーライド行動が生じていないことを示している。これに対し、我々の研究は患者一人あたりのHLA適合ドナーの数が少ないという情報が20代男性の提供意向を高めていることを明らかにしており、これは彼らがHLA適合ドナーの数を過度に多く見積もり、フリーライド行動を取っていることを示唆するものである。

本論文の構成は以下の通りである。第2節では、JMDPにおけるコーディネーションの過程を概観し、フィールド実験の詳細を説明する。第3節で結果を示し、第4節で議論と結論を述べる。

# 2 Field Experiment

## 2.1 Background: Coordination Process of JMDP

フィールド実験の介入のタイミングの理解を促すために、JMDPに登録した潜在ドナーによる幹細胞の提供までのコーディネーション過程を概観しておく。はじめに、潜在ドナーがJMDPに登録した患者とマッチすると、JMDP事務局は彼らに幹細胞提供を依頼する適合通知（compatibility notice）を送付する[[5]](#footnote-25)。適合ドナーは問診票を記入し、提供意思の有無を示して、適合通知に返信する。

その後、移植に向けたコーディネーションが始まる。適合ドナーはおよそ1カ月以内に確認検査（confirmatory typing）を受ける。この段階では、コーディネーターは提供手順の詳細（骨髄採取もしくは末梢血幹細胞採取）を説明し、適合ドナーとその家族に提供の意向を調査する。潜在ドナーは二つの採取方法を選べる。加えて、コーディネーター医師が問診、検診、感染症の有無と血液型を調べる採血をする。これらのテストは適合ドナーがJMDPの定める基準に満たしているかどうかを検査する。

患者は同時に最大10人の適合ドナーとのコーディネーションを受けられる。患者の医師は確認検査を受けた適合ドナーから最も適切な候補者を選ぶ。重要なこととして、適合ドナーはマッチしている患者の情報（他の適合ドナーと何人マッチしているかなど）を一切知ることができない。また、適合ドナーがコーディネーターやコーディネーター医師からそのような情報を得ることもできない。

最も適切な候補者として選ばれた適合ドナーはコーディネーターやコーディネーター医師からの説明を受けた後に、最終同意をする必要がある。同時に、ドナーの家族の代表者も採取に同意しなければならない。ただし、これ以降、選定ドナーは意向を変えられない。最終同意後、選定ドナーは術前検査や採取の準備を受けるために1週間程度の入院をする。この後、ドナーは幹細胞採取の手術を受ける。確認検査から採取までの期間はおよそ3–4カ月である。

## 2.2 Experimental Design



Figure 2.1: Intervention Messages

我々の実験はJMDPが適合ドナーに幹細胞提供を依頼する適合通知の内容に介入を施す。図 2.1 は我々の介入を示す。適合通知は提供ドナーが7日以内に返信するべきことを注記している。また、JMDPは適合通知と一緒に前小節で概観したコーディネーション過程を記述したハンドブックを同封している。

我々はコーディネーションを促進するために適合通知に二つのメッセージを加えた[[6]](#footnote-31)。Probability messageは登録患者一人あたりの適合ドナーの数が少ないことを強調している。第1節で述べたように、同じHLA型を持つ他の潜在ドナーがプールに存在すれば、自身の提供はその適合ドナーと代替可能なので、移植は公共財の性質を有している。したがって、ボランティアのジレンマの予測通り、一般的なHLAの型を持っている人ほど提供に消極的となるだろう。

しかしながら、幹細胞移植の文脈では、適合ドナーがグループサイズを正確に知ることができないので、グループサイズの期待値が意思決定に影響を与える。すなわち、同じHLA型を持つ適合ドナーの数が多いと予測しているほど、そのドナーは提供に消極的となるだろう。もし、あるドナーの信念が過度に高いならば、Probability messageは信念の下方修正を通じて、フリーライド行動を抑制し、提供意向を高めるだろう。ただし、逆の効果も考えられる。あるドナーの信念が真のグループサイズよりも過小ならば、Probability messageは信念の上方修正を通じて、フリーライド行動を誘発し、提供意向を下げる恐れがある。

Early Coordinationメッセージは早いコーディネーションが患者の移植率を高めることを提示している。適合通知への返信が遅れることによってコーディネーション期間が長くなると、患者の移植率が減少してしまう。したがって、このメッセージは患者への移植によって得られる効用が経時的に減衰していくことを暗に意味している。第1節で述べたように、適合通知への返信は行動経済学における先延ばし行動が障害となりうる。もしあるドナーが患者への移植によって得られる効用水準が時間に対して一定であると信じているならば、今コストを支払ってコーディネーションを受けるよりも将来コストを支払ってコーディネーションを受けようと考えるはずである。そして、同じ意思決定が将来でも生じ、適合通知への返信が遅れてしまう可能性がある。このメッセージはこの先延ばし行動を抑制することを目的としている。仮にこのメッセージが全体の適合通知への返信率を高めなくても、早期返信を促進しているだろう。

我々は二つの介入メッセージの効果を推定するために四つの実験群を設けた。実験群Aは介入メッセージなしの標準的な適合通知を送付した（コントロール群）。実験群BとCはそれぞれ確率メッセージとEarly Coordinationメッセージを追加した適合通知を送付した。実験群Dは二つの介入メッセージを同時に加えた適合通知を送付した。この実験群は情報過多による認知負荷の負の影響を検証する目的で設けた。

Table 2.1: Assignment Schedule

| week | Sep 21 | Oct 21 | Nov 21 | Dec 21 | Jan 22 | Feb 22 |
| --- | --- | --- | --- | --- | --- | --- |
| Week 1 | B | C | C | D | B | A |
| Week 2 | D | B | A | A | C | B |
| Week 3 | A | D | B | C | D | C |
| Week 4 | C | A | D | B | A | D |

実験参加者は2021年9月から2022年2月の間に適合通知を受け取った11,154名の適合ドナーである。JMDP事務局に無理のない範囲でランダムネスを保つために、我々は実験群を週単位のクラスターランダム化で割り当てた。このとき、可能な限り週と月で実験群がバランスするように設計している。表 2.1 に割り当てスケジュールをまとめている。本実験の実施にあたって、我々はメッセージの内容を含めて大阪大学経済学研究科とJMDPの倫理審査の承認を得ている。

## 2.3 Data and Empirical Strategy

我々はJMDPの協力のもと、コーディネーションデータを2022年6月末に受け取った。観測単位は実験参加者である。個人属性として、データは性別・年齢・コーディネーション経験回数・都道府県レベルの在住地区を記録している。コーディネーション過程として、データは各段階（適合通知への返信・確認検査・候補者選定・最終同意・採取）に到達したかどうかを記録している。我々はこれらの変数をアウトカム変数として用いる。加えて、適合通知への返信について、データは返信日数と提供意向を記録している。さらに、コーディネーションが中断したならば、データは中断理由を三つのカテゴリ（患者側の理由・健康以外のドナー理由・ドナーの健康理由）で記録している。分析では、日本在住かつコーディネーションが（中断を含めて）終了している11,049名の適合ドナーを用いる[[7]](#footnote-34)。

追加的なデータとして、我々はJMDPがウェブ上で公開している医療機関リストを用いる[[8]](#footnote-35)。このリストは完全な住所・骨髄採取（BM collection）が可能かどうか・末梢血幹細胞採取（PBSC collection）が可能かどうかを含んでいる。我々はこのリストを都道府県レベルで集約し、10平方キロメートルあたりの病院数を計算し、それをコーディネーションデータと都道府県をマージキーにして突合した。我々はこの変数をコーディネーションや提供のトラベルコストとして考慮する。

Table 2.2: Overview of Field Experiment

|  | Experimental Arms | | | |  |
| --- | --- | --- | --- | --- | --- |
|  | A | B | C | D | F-test, p-value |
| A. Interventions | | | | | |
| Standard notification | X | X | X | X |  |
| Probability message |  | X |  | X |  |
| Early Coordination message |  |  | X | X |  |
| B. Sample Size | | | | | |
| N | 2535 | 3053 | 2726 | 2735 |  |
| C. Balance Test | | | | | |
| Male (=1) | 0.624 | 0.633 | 0.631 | 0.609 | 0.424 |
| Age | 38.376 | 38.121 | 37.448 | 37.978 | 0.096 |
| Number of past coordinations | 1.609 | 1.589 | 1.625 | 1.563 | 0.515 |
| Number of listed hospitals | 0.476 | 0.490 | 0.487 | 0.485 | 0.826 |
| Number of hospitals listed with PBSC collection | 0.162 | 0.167 | 0.166 | 0.164 | 0.831 |
| Number of hospitals listed with BM collection | 0.246 | 0.256 | 0.254 | 0.251 | 0.711 |
| Notes: Balance test regresses a covariate on treatment dummies and test a null hypothesis that all coefficients are zero. We use the robust standard errors for statistical inference. | | | | | |

表 2.2 はフィールド実験の概要である。パネルAは各実験群の介入を示しており、パネルBは各実験群のサンプルサイズを示している。パネルCはランダム化が成功しているかどうかを検証するバランステストである。年齢を除く変数について実験群間で平均的に差がないので、実験群の割り当てはおおよそランダムである。ただし、年齢は群間でばらつきがある。実験群Cの平均年齢はコントロール群と比較して1歳程度若い。

我々はどのような情報を潜在ドナー全員に与えるべきかに関心があるので、母集団全体の平均処置効果を識別したい。先決変数で条件づけたとき、実験群の割り当ては潜在成果変数（ある実験群が割り当てられたときに観察されるであろう成果変数）と独立であるはずなので、先決変数で条件づけた実験群間の平均値の差で平均処置効果を識別できる。したがって、我々は月の週に適合通知を受け取った個人についての以下の線形確率モデルを推定する。

ここでは個人属性ベクトルである。ある時期の共通のショックをコントロールするために、月および週のダミー変数とを加える。我々の関心のあるパラメータはである。固定効果を加えたとき、観察不可能な要素のクラスター（実験週）内の相関を生み出す原因はないと思われる。したがって、我々は推定値の統計的推論の際に頑健標準誤差を用いる[[9]](#footnote-38)。

# 3 Experimental Results

## 3.1 Effects on Reply and Intention

初めに、適合ドナーの意向を最も反映していると考えられる返信への介入効果を推定する。返信のアウトカム変数は提供の意向に関わらず適合通知に返信したならば1を取るダミー変数である。コントロール群（実験群A）の返信率は%である。

我々は返信へのメッセージ効果を二つに分解する。第一に、提供意向を伴った返信への効果である。アウトカム変数は潜在ドナーが適合通知に返信し、かつ提供意思を示したならば1を取るダミー変数である。第二に、提供意向を伴わない返信に対する効果である。アウトカム変数は潜在ドナーが適合通知に返信し、かつ提供意思を示していないならば1を取るダミー変数である。ただし、これらの効果を推定するとき、我々は未返信者のアウトカム変数を0とコーディングして、サンプルに含める。したがって、二つのアウトカム変数の和は必ず返信のダミー変数となるので、正の意向と負の意向に対する効果の和は返信に対する効果となる。コントロール群では、正の意向を伴う返信率は%である一方で、負の意向を伴う返信率は%であった。したがって、返信者の%は提供に積極的である。

適合ドナーが返信する前に、患者側の都合でコーディネーションが中断してしまう可能性がある。これは潜在ドナーの意向と独立に生じるので、我々は各実験群でそのようなケースに該当する–%を分析サンプルから除外している。ただし、このケースは実験群間でバランスしている（F-value, p-value = 0.643）。

Table 3.1: Linear Probability Model of Reply and Intention

|  | | | Intention | | | |
| --- | --- | --- | --- | --- | --- | --- |
|  | Reply | | Positive | | Negative | |
|  | (1) | (2) | (3) | (4) | (5) | (6) |
| Treatment B | 0.0105 | 0.0139\*\* | 0.0218 | 0.0210 | -0.0113 | -0.0071 |
|  | (0.0075) | (0.0059) | (0.0135) | (0.0123) | (0.0094) | (0.0076) |
| Treatment C | -0.0057 | 0.0026 | -0.0053 | -0.0029 | -0.0004 | 0.0054 |
|  | (0.0081) | (0.0047) | (0.0151) | (0.0112) | (0.0150) | (0.0094) |
| Treatment D | 0.0058 | 0.0062 | 0.0046 | 0.0058 | 0.0012 | 0.0004 |
|  | (0.0074) | (0.0052) | (0.0166) | (0.0101) | (0.0138) | (0.0069) |
| Control average | 0.8835 | 0.8835 | 0.5533 | 0.5533 | 0.3303 | 0.3303 |
| Covariates |  | X |  | X |  | X |
| Num.Obs. | 10985 | 10985 | 10985 | 10985 | 10985 | 10985 |
| Notes: \* p < 0.1, \*\* p < 0.05, \*\*\* p < 0.01. The robust standard errors are reported in parentheses. Covariates are gender, squared polynomial of (demeaned) age, number of past coordinations, number of hospitals per 10 square kilometers, number of hospitals with PBSC collection per 10 square kilometers, number of hospitals with BM collection per 10 square kilometers, month dummies, and week dummies. | | | | | | |

表 3.1 は線形確率モデルの推定結果である。奇数列は実験群ダミーだけを説明変数に入れており、単純な平均値の差に対応している。偶数列は実験群ダミーに加えて、個人属性・月および週の固定効果をコントロールしている。推定結果はすべての実験群が、共変量の有無に関わらず、返信や意向に対して統計的に有意な効果を持っていない。我々は代替手法としてロジットモデルを推定（表 4.1 ）し、同様の結果を得ている（オッズ比の95%信頼区間が1を含んでいる）。

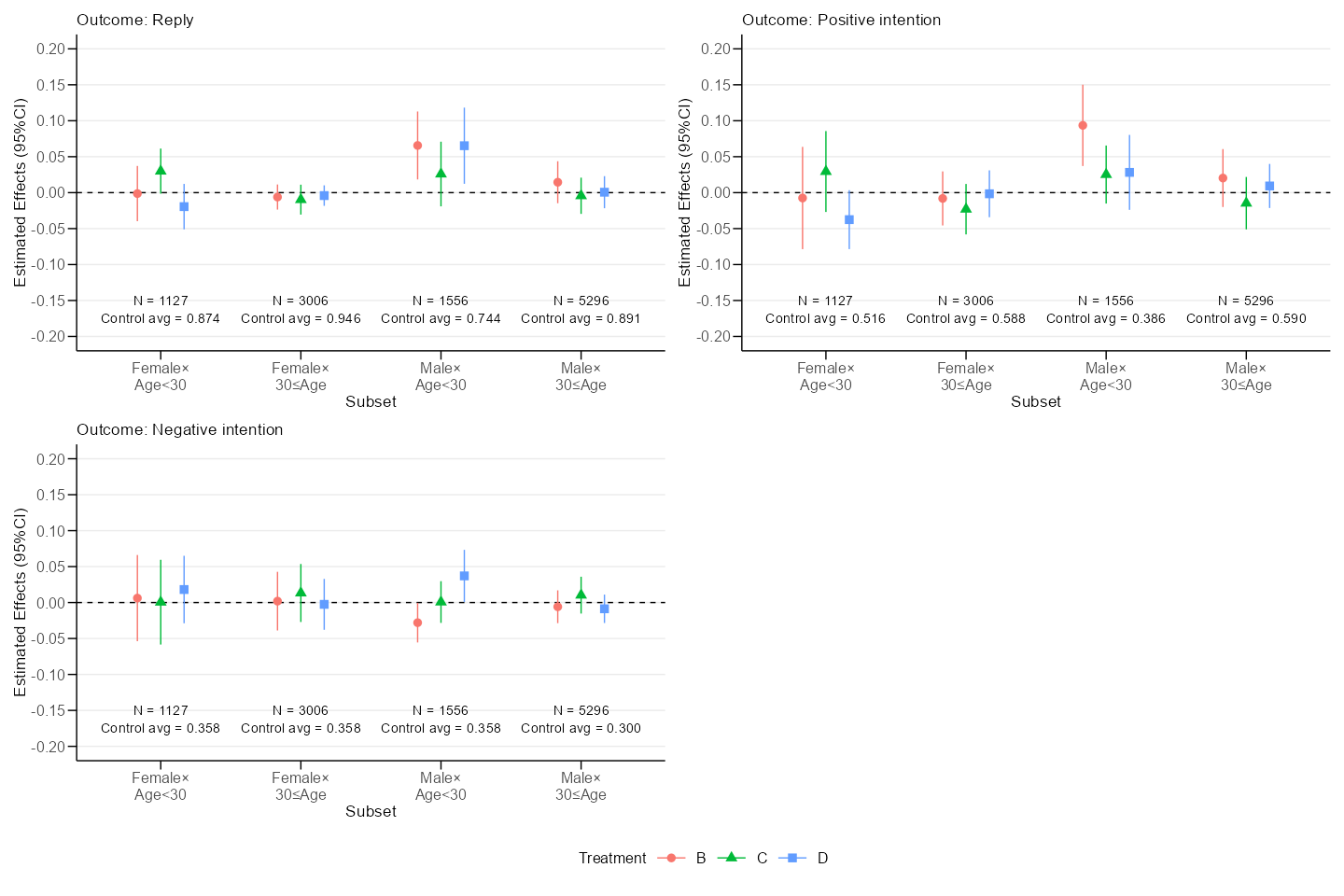


Figure 3.1: Effect on Reply and Intentions by Gender and Age Group. Note: These plots show the average effect (and associated 95% confidential interval) on each outcome by gender and age group. We use robust standard errors We control number of past coordinations, number of hospitals per 10 square kilometers, number of hospitals with PBSC collection per 10 square kilometers, number of hospitals with BM collection per 10 square kilometers, month dummies, and week dummies.

次に、メッセージ効果の異質性を検証するために、我々はサンプルを性別と年齢層（30歳未満か否か）で4つのサブセットに分割し、各サブセットでメッセージ効果を推定する。図 3.1 は係数プロットである。結果として、確率メッセージを含む実験群BとDが20代男性の返信率を約6ポイント（コントロール平均が%なので、%増加）高めており、これは統計的に5%水準で有意である。特に、確率メッセージのみを加えた実験群Bが正の意向を伴う返信を約10ポイント（コントロール平均が%なので、%増加）増やしており、これも統計的に1%水準で有意である（表4.3）。ただし、他の性・年代グループで、我々の介入は返信や意向に対して統計的に有意な効果をもたない。

コントロール群において、適合通知に返信した20代男性の約%が提供に積極的であり、これは他の性・年代グループの返信者の意向比率（60%）よりも低い[[10]](#footnote-46)。実験群B（確率メッセージ）は20代男性の返信者の意向比率を50%から%に引き上げている。さらに、若い男性の移植成績は他の性・年代よりも良いとされている。これらを総合すると、ドナーとして望ましい人の行動変容（コーディネーションの開始）を促しているという意味で、確率メッセージはコーディネーションの効率性を改善しているといえよう。

## 3.2 Random Causal Forest

介入効果の異質性を深堀りするために、我々は完全にノンパラメトリックに介入効果の異質性を検証できるrandom causal forest (RCF)を用いる(Athey and Imbens 2016; Wager and Athey 2018)。この手法は回帰木アルゴリズムに基づいており、観察可能な特徴で条件づけた平均処置効果を推定する。この手法は労働(Davis and Heller 2017)・教育(Carlana, La Ferrara, and Pinotti 2022)・節電(Murakami et al. 2022)などの広い文脈で用いられている。

このアルゴリズムは研究者によって与えられた共変量の一つを用いて、サンプルを二つのサブサンプル（リーフ）に分割する。具体的に、このアルゴリズムはであるサンプルとであるサンプルに分割する。回帰木はアウトカム変数の平均二乗誤差（誤差はあるリーフに属するサンプルのアウトカムの平均値と観察されるアウトカムの差である）を最小にするように特定の閾値を決める。RCFはあるリーフに属する平均処置効果（条件付き平均処置効果）を推定することが目的なので、予測処置効果の期待平均二乗誤差を最小にするように閾値を決める。この最小化はリーフ間の条件付き平均効果の分散（異質性）を高めかつ、リーフ内の分散を抑えることで達成される。これを終端条件に達するまで各リーフで二つのリーフに分けるプロセスを繰り返す。RCFは終端のリーフにおける条件付き平均処置効果を予測する。

回帰木アルゴリズムの欠点は予測値の分散が大きく（過学習）なり、予測精度が落ちることである。これを防ぐために、RCFはアンサンブル法と呼ばれるアルゴリズムを導入している。これは全データから重複なしのサブサンプルを数千個作り、各サブサンプルで回帰木を育てる。最終的な結果は数千個の回帰木の予測値の平均とする[[11]](#footnote-48)。

大雑把に言えば、RCFは観察される個人属性で介入の処置効果を予測する。すなわち、介入を受けていない人についても、その人の属性に基づいて処置効果を予測できる。ただし、この手法は処置の割り当てが潜在成果変数と条件つきにランダムであることを仮定しなければならない。実験群はおおよそランダムに割り当てられているので、我々はこの手法を用いることができる。この小節では、意向を伴う返信への効果に着目する。RCFに用いる共変量は性別・年齢・過去のコーディネーション経験・在住地域の病院の密度である。

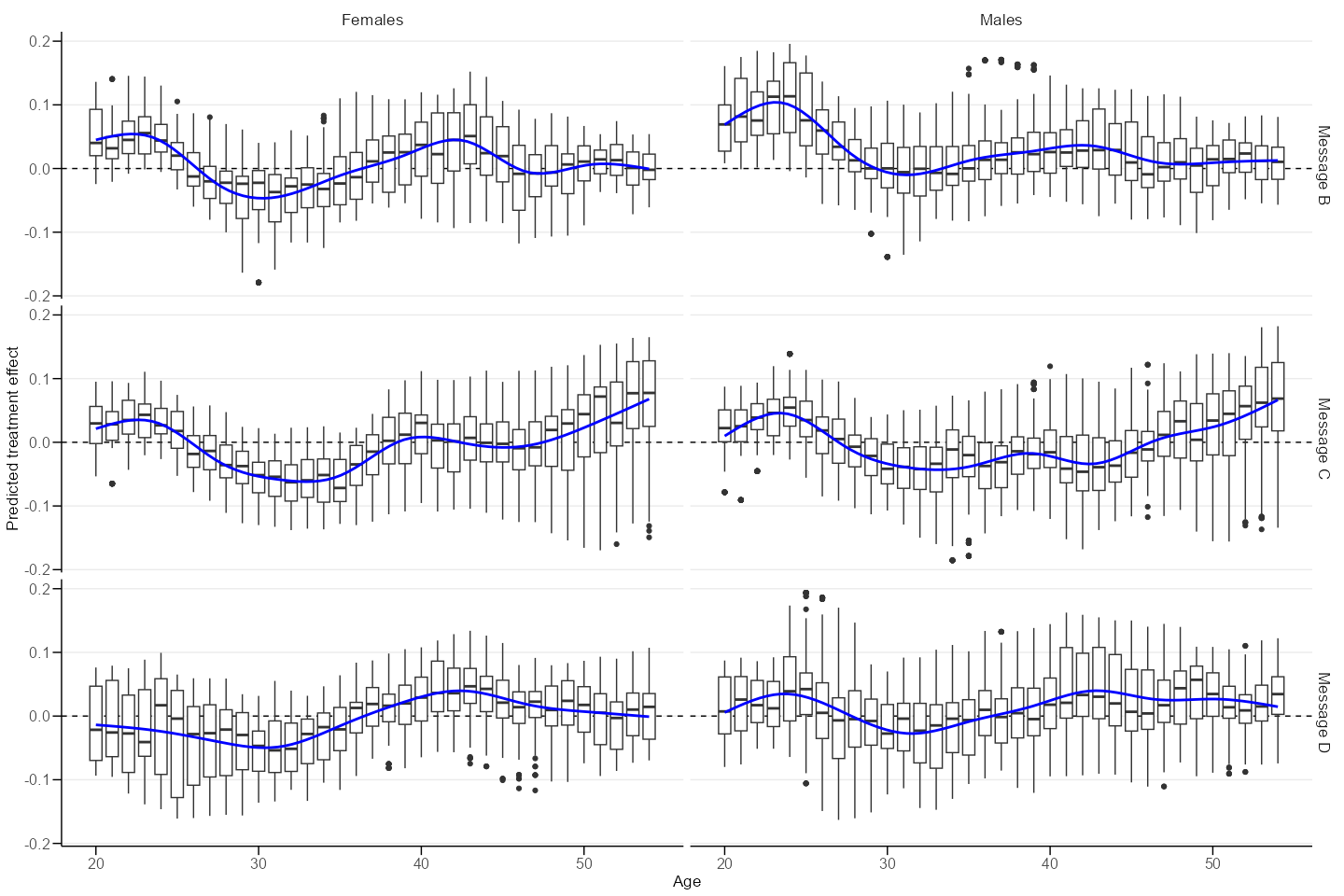


Figure 3.2: Boxplot of Predicted Treatment Effects by Gender and Age. Notes: Blue fitted line represents GAM smoothing. Covariates are gender, age, number of past coordinations, number of hospitals per 10 square kilometers, number of hospitals with PBSC collection per 10 square kilometers and number of hospitals with BM collection per 10 square kilometers.

図3.2は意向を伴う返信に対する各実験群の予測トリートメント効果の分布を年齢別にプロットしている。注目するべきこととして、25歳以下の男性のすべてに対して、実験群Bの効果は正である。結果として、RCFは20代男性における実験群Bの平均効果は12.1ポイントであり、図3.1で示した線形確率モデルのサブサンプル分析の結果と近い値である（表4.5参照）。また、30–40代男性の一部に対して、実験群Bの効果は10ポイント以上である。そのような男性は相対的に病院の数が多い地域に住んでおり、コーディネーションや提供のトラベリングコストが低い。

Early Coordinationメッセージのみを加えた実験群Cの効果は大きく異質的である。30代前半の女性に対して、ほとんどのトリートメント効果は負である。40代後半以上の男女に対して、処置効果の中央値は正であるが広く分布している。驚くべきことに、この年齢層の男女において、実験群Cが正の効果を持つ人は相対的に病院の数が少ない（トラベリングコストが高い）地域に住んでいる。

Table 3.2: Correlation of Predicted Treatment Effects

|  | Treatment effect D | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Females | | Males | | Females | | Males | |
|  | Age<30 | 30≤Age | Age<30 | 30≤Age | Age<30 | 30≤Age | Age<30 | 30≤Age |
|  | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Constant | -0.0351\*\*\* | 0.0077\*\*\* | -0.0122\*\*\* | 0.0133\*\*\* | -0.0310\*\*\* | 0.0072\*\*\* | 0.0080\*\*\* | 0.0086\*\*\* |
|  | (0.0019) | (0.0007) | (0.0022) | (7e-04) | (0.0018) | (0.0008) | (0.0021) | (0.0007) |
|  | [-0.0387, -0.0314] | [0.0062, 0.0091] | [-0.0166, -0.0078] | [0.0120, 0.0146] | [-0.0344, -0.0276] | [0.0057, 0.0087] | [0.0039, 0.0121] | [0.0073, 0.0099] |
| B + C | 0.3107\*\*\* | 0.3559\*\*\* | 0.3337\*\*\* | 0.4022\*\*\* |  |  |  |  |
|  | (0.0226) | (0.0073) | (0.0192) | (7e-03) |  |  |  |  |
|  | [0.2663, 0.3552] | [0.3416, 0.3702] | [0.2960, 0.3713] | [0.3884, 0.4159] |  |  |  |  |
| B |  |  |  |  | -0.0529 | 0.3994\*\*\* | -0.1354\*\*\* | 0.6256\*\*\* |
|  |  |  |  |  | (0.0377) | (0.0145) | (0.0261) | (0.0157) |
|  |  |  |  |  | [-0.1268, 0.0211] | [0.3710, 0.4279] | [-0.1867, -0.0842] | [0.5948, 0.6565] |
| C |  |  |  |  | 0.7489\*\*\* | 0.3195\*\*\* | 0.9609\*\*\* | 0.2693\*\*\* |
|  |  |  |  |  | (0.0377) | (0.0125) | (0.0307) | (0.0103) |
|  |  |  |  |  | [0.6750, 0.8228] | [0.2950, 0.3441] | [0.9007, 1.0210] | [0.2490, 0.2896] |
| Num.Obs. | 1127 | 3006 | 1556 | 5296 | 1127 | 3006 | 1556 | 5296 |
| Notes: \* p < 0.1, \*\* p < 0.05, \*\*\* p < 0.01. Robust standard errors are reported in parentheses. Square brackets show 95 percent confidential intervals. | | | | | | | | |

実験群Dは確率メッセージ（実験群B）とEarly Coordinationメッセージ（実験群C）の両方を加えている。したがって、もし適合ドナーが実験群Dの情報を完全に理解しているならば、実験群Dの効果は実験群BとCの効果の合計となるはずである。この点を検証するために、表3.2の列(1)–(4)は実験群Dの予測処置効果を実験群BとCの予測効果の合計で回帰している。どの性・年代でも実験群BとCの効果の合計が1ポイント増えると、実験群Dの効果は0.3–0.4ポイント増加する。95%信頼区間は1を含まないので、この増分は1ポイントより統計的に有意に低い。この結果は適合ドナーが実験群Dの情報を完全に理解しておらず、情報過多による認知負荷になっていること示唆している。

完全に情報を受け取っていない適合ドナーは確率メッセージもしくはEarly Coordinationメッセージのどちらかを重視するか、二つの情報を同程度割り引いて評価しているだろう。この点を検証するために、表3.2の列(4)–(8)は実験群Dの予測処置効果を実験群Bの予測効果と実験群Cの予測効果で回帰している。その結果、20代男女では、実験群CとDの偏相関は実験群BとDの偏相関よりも強い。すなわち、彼らは確率メッセージよりもEarly Coordinationメッセージを重視している。したがって、確率メッセージを提供しているにも関わらず、実験群Dは20代男性に効果がないのである。

興味深いことに、30歳以上の男女では、実験群CとDの偏相関は実験群BとDの偏相関よりも弱い。すなわち、彼らはEarly Coordinationメッセージよりも確率メッセージを重視している。したがって、40代後半以上の男女における実験群Cの効果の中央値が8ポイント程度であるにも関わらず、実験群Dの処置効果の中央値はほとんどゼロである。

## 3.3 Response Speed to Notification

Early Coordinationメッセージは早いコーディネーションが患者の移植率を高めるという事実を提供している。したがって、このメッセージは適合通知への早期返信を促している可能性がある。この可能性を検証するために、我々は特定の期間以内に返信する確率への効果を線形確率モデルで推定する。

アウトカム変数は潜在ドナーが日以内に返信したならば1を取るダミー変数である。日以降に返信した・返信しなかった場合、アウトカム変数は0となる。の値を十分に大きくすれば、この分析は第3.1節の結果と一致する。さらに、第3.1節と同様に、我々はドナーの意向の視点から二つの効果に分解する。一つは提供意向を伴う日以内の返信への効果である。アウトカムは日以内に返信し、かつ提供の意向を示したならば1を取るダミー変数である。もう一つは提供意向を伴わない日以内の返信への効果である。アウトカムは日以内に返信し、かつ提供の意向を示さなければ1を取るダミー変数である。日以降に返信した・返信しなかった場合、これらのアウトカムも0とコーディングされる。我々はに焦点を当てる。

コントロール群では、返信者の平均的な返信日数は10日である。また、正もしくは負の意向を伴う返信の平均返信日数はそれぞれ9日・12日である。最長の返信日数は46日である。加えて、適合通知に目安として記載されている7日以内の返信は33%であり、正の意向を伴う7日以内の返信は23%である。

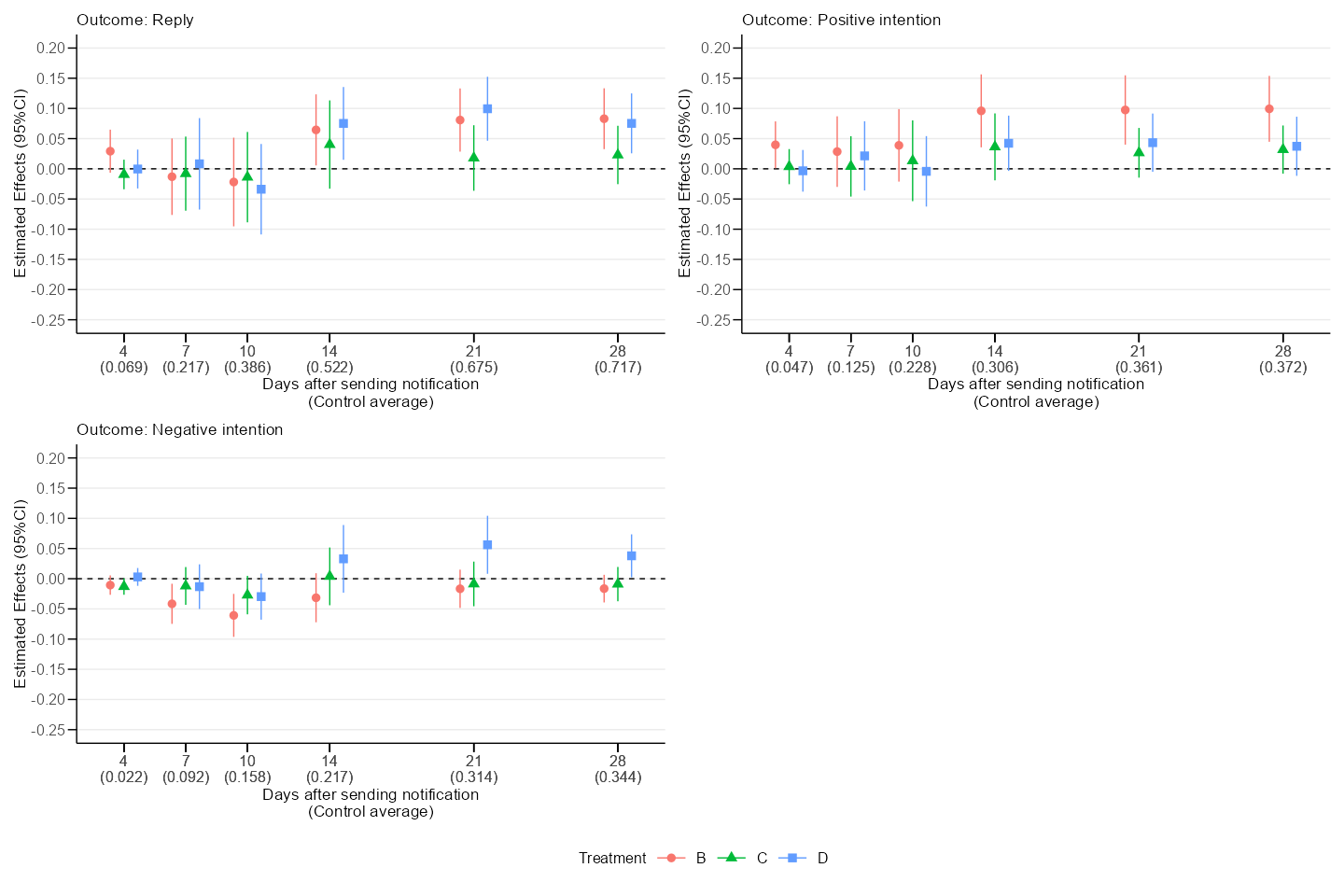


Figure 3.3: Effect on Reply within Specific Days after Sending Notification among Males Less than 30. Notes: These plots show the average effect (and associated 95% confidential interval) on each outcome. We use robust standard errors We control number of past coordinations, number of hospitals per 10 square kilometers, number of hospitals with PBSC collection per 10 square kilometers, number of hospitals with BM collection per 10 square kilometers, month dummies, and week dummies.

先の小節で見たように返信や意向への効果は異質的であるので、この小節は性・年代でサブサンプルに分割して、返信日数への効果の異質性に注目する[[12]](#footnote-59)。図3.3は20代男性の結果である。このサブグループでは、我々は確率メッセージのみを加えた実験群Bが提供意向を伴う全体の返信率を10ポイント程度高めていたことをすでに明らかにしている（図3.1）。図3.3はその効果が14日以内の返信から生じていることを示している。実験群Bは提供意向を伴う10日以内の返信を（統計的に有意ではないが）5ポイント程度高めており、14日以内の返信を10ポイント程度高めている。すなわち、実験群Bは提供意向を伴う10日以内の返信を5ポイント、提供意向を伴う10–14日以内の返信を5ポイント高め、これらが全体の返信率を10ポイント高めていることにつながっている。また、このグループでは、早期返信を促している可能性のある実験群C（Early Coordinationメッセージ）が短期間の返信を高めているということはなかった。

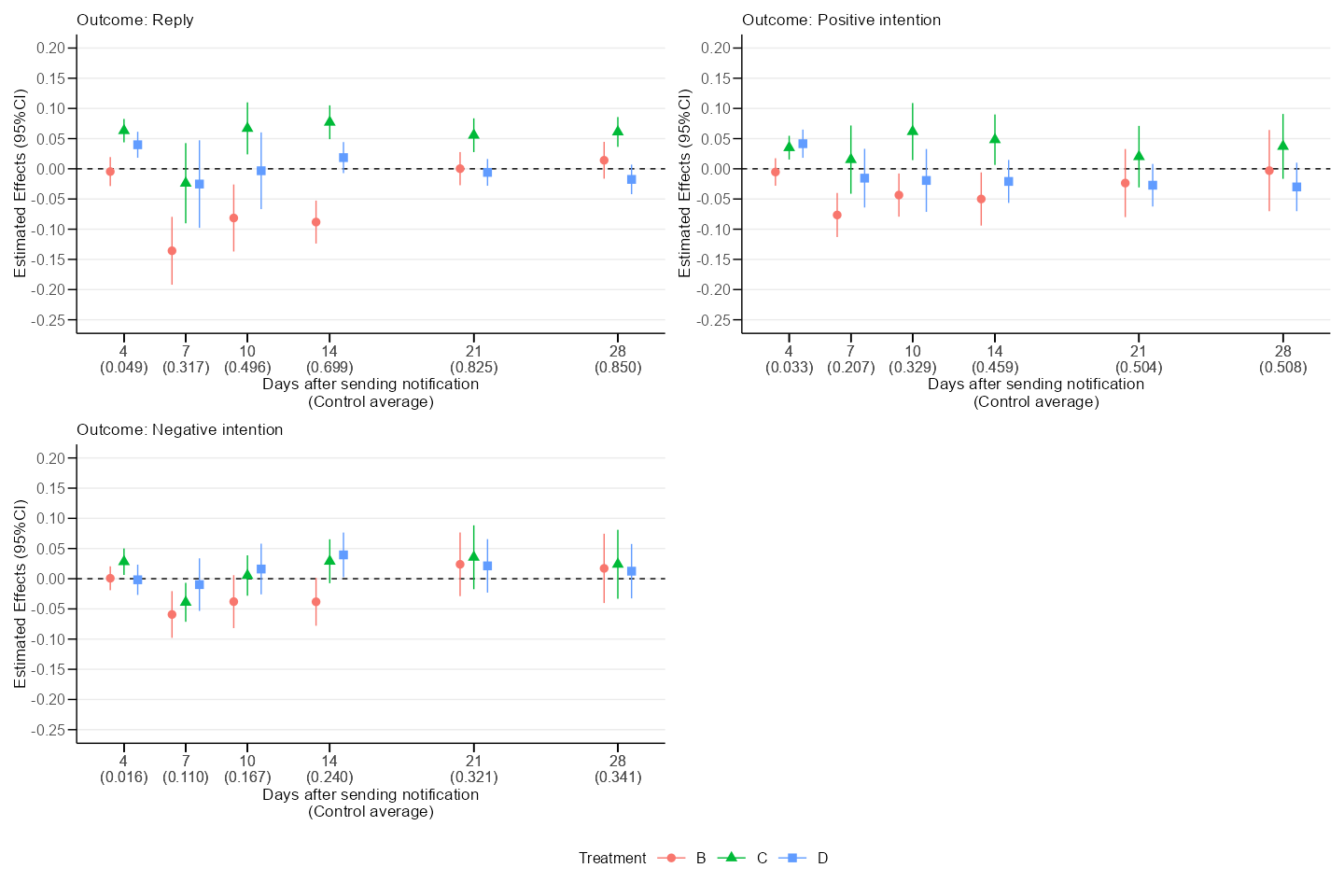


Figure 3.4: Effect on Reply within Specific Days after Sending Notification among Females Less than 30. Notes: These plots show the average effect (and associated 95% confidential interval) on each outcome. We use robust standard errors We control number of past coordinations, number of hospitals per 10 square kilometers, number of hospitals with PBSC collection per 10 square kilometers, number of hospitals with BM collection per 10 square kilometers, month dummies, and week dummies.

しかしながら、実験群Cは20代女性の早期返信を促している可能性がある。図3.3は20代女性の結果である。図3.1より、実験群Cは20代女性の全体の返信率に統計的に有意な影響を与えていない。一方で、この実験群は4日以内の返信率を6ポイント程度高めており、これは5%水準で有意である。この効果は28日以内の返信率でも維持している。すなわち、4–28日の返信率の変動はコントロール群と比較して同じである。そして、実験群Cの全体の返信率に対する点推定は3ポイントなので、この実験群は28日以降の返信率を3ポイント下げている。したがって、実験群Cは4日以内の返信率を高め、それ以降の返信率を若干減らしている。言い換えれば、Early Coordinationメッセージは20代女性の早期返信を促している。4日以内の返信の増加は正の意向を伴う返信と負の意向を伴う返信を同程度高めていることから生じている。

また、Early Coordinationメッセージを含む実験群Dの返信への効果も実験群Cと同様の傾向である。ただし、実験群Cの結果と異なり、4日以内の返信の増加は正の意向を伴う返信のみの増加から生じている。

## 3.4 Effects on the Coordination Process

最後に、我々は適合通知への返信以降のコーディネーション過程の各工程への影響を検証する。第2.1節で説明したように、コーディネーション過程は確認検査・候補者選定・最終同意・提供の4つのステージからなる。我々は適合ドナーが各ステージに到達したならば1を取るダミー変数をアウトカム変数として用いる。

第3.1節の分析と同様に、我々は適合ドナーの意思とは無関係にコーディネーションが中断したと思われるサンプルを除く。確認検査への効果を推定するとき、我々は患者側の理由で中断した事例を除外する。患者の意思は候補者選定時に最も健康な適合ドナーを選ぶ可能性が高いので、候補者選定以降でドナーの健康理由で中断した事例はドナーの意思と無関係に生じていると考えられる。よって、候補者選定・最終同意・提供に対する効果を推定するとき、我々は患者側の理由で中断したサンプルと候補者選定以降にドナーの健康理由で中断したサンプルを除く。ただし、サンプルの除外だけでは医者の意思決定を完全に取り除けないかもしれないので、解釈には注意を要する。コントロール群では、適合ドナーの24%が確認検査を受けて、8%が候補者になり、6%が最終的に提供した。

Table 3.3: Linear Probability Model of Coordination Process

|  | CT | | Candidate | | Consent | | Donation | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Treatment B | 0.0325\*\* | 0.0326\*\*\* | 0.0051 | 0.0033 | 0.0059 | 0.0037 | 0.0040 | 0.0020 |
|  | (0.0141) | (0.0073) | (0.0076) | (0.0050) | (0.0074) | (0.0041) | (0.0060) | (0.0043) |
| Treatment C | 0.0146 | 0.0136 | 0.0010 | -0.0023 | 0.0024 | -0.0011 | 0.0016 | -0.0017 |
|  | (0.0163) | (0.0087) | (0.0102) | (0.0057) | (0.0086) | (0.0044) | (0.0076) | (0.0047) |
| Treatment D | 0.0260 | 0.0299\*\*\* | 0.0084 | 0.0087\* | 0.0099 | 0.0098\* | 0.0030 | 0.0029 |
|  | (0.0161) | (0.0077) | (0.0093) | (0.0046) | (0.0102) | (0.0048) | (0.0095) | (0.0053) |
| Control average | 0.2350 | 0.2350 | 0.0779 | 0.0779 | 0.0687 | 0.0687 | 0.0574 | 0.0574 |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
| Covariates |  | X |  | X |  | X |  | X |
| Num.Obs. | 10435 | 10435 | 8587 | 8587 | 8558 | 8558 | 8441 | 8441 |
| Notes: \* p < 0.1, \*\* p < 0.05, \*\*\* p < 0.01. The robust standard errors are reported in parentheses. Covariates are gender, squared polynomial of (demeaned) age, number of past coordinations, number of hospitals per 10 square kilometers, number of hospitals with PBSC collection per 10 square kilometers, number of hospitals with BM collection per 10 square kilometers, month dummies, and week dummies. | | | | | | | | |

表3.3はフルサンプルの推定結果を示している。その結果、確率メッセージを含む実験群BとDは確認検査の実施率を約3ポイント高めており、これは統計的に有意な効果である。この効果は実験群BとDの返信への効果よりも大きい。これは実験群BとDでは、健康理由を含むドナー側の都合でコーディネーションが中断する可能性がコントロール群と比較して小さいからである。したがって、実験群BとDは全体的に提供に積極的な人を増やしていないが、提供の意向を強固にしており、コーディネートの継続性に貢献していると考えられる。

実験群BとDの候補者選定以降のステージに対する効果は1ポイント以下であり、統計的に非有意である。これは解釈に注意が必要である。我々の介入は需要（患者数）に影響を与えていないので、我々の介入によって確認検査に到達する人が増えたならば、その分だけ外生的な理由で候補者に選定されなかった人を増やしているはずである。実験群Bとコントロール群で、確認検査を実施した潜在ドナーはそれぞれ774人と564人である。すなわち、実験群Bは確認検査実施数を210件増やしたことになる。また、患者理由もしくはドナーの健康理由で候補者に選定されなかった人は実験群Bとコントロール群でそれぞれ556人と385人である。すなわち、実験群Bは外生的な理由で候補者に選定されなかった人を171人増やしている。したがって、推定値は我々の介入効果だけでなく幹細胞移植の需要の影響を含んでいる。

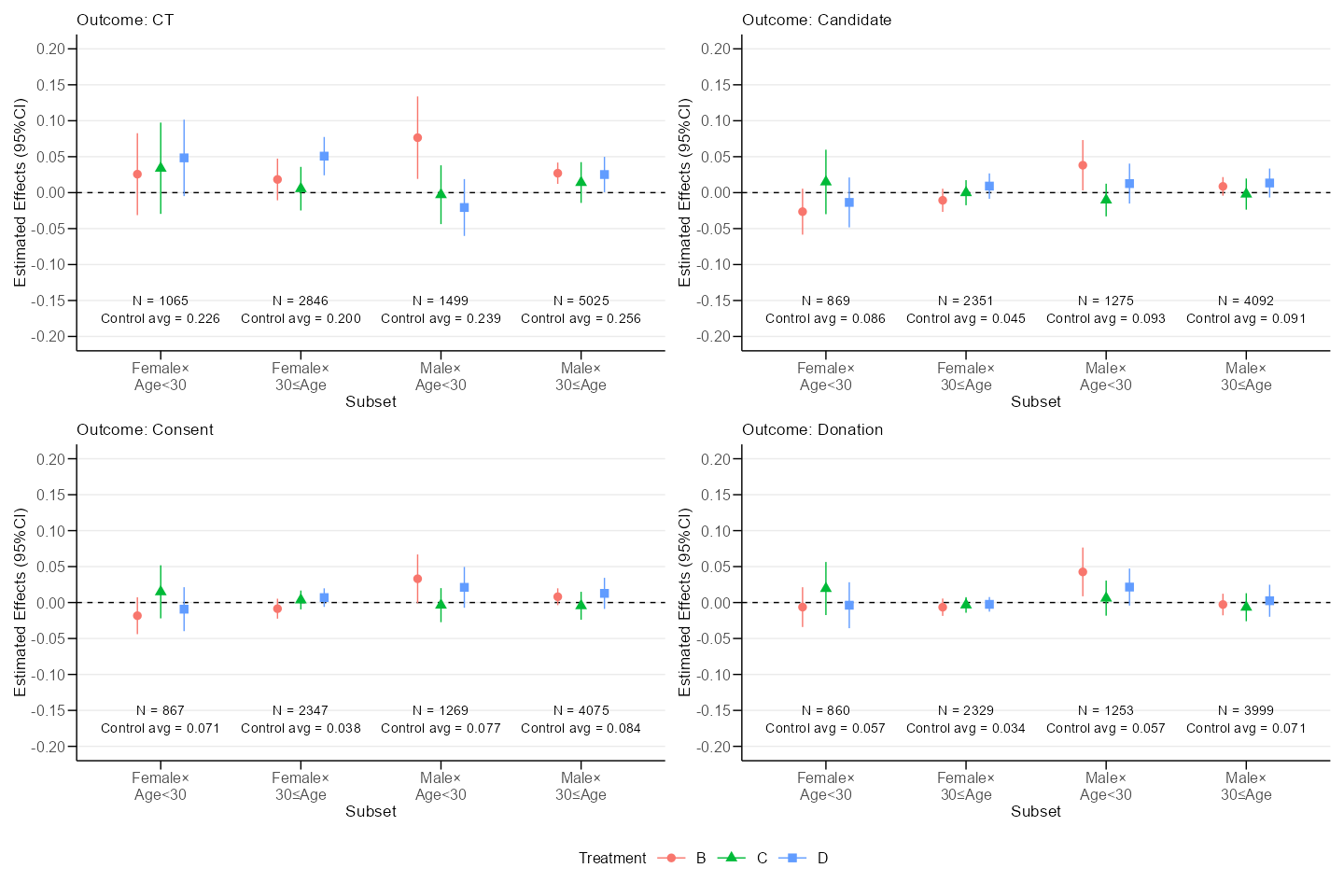


Figure 3.5: Effect on Coordination by Gender and Age Group. Note: These plots show the average effect (and associated 95% confidential interval) on each outcome by gender and age group. We use robust standard errors We control number of past coordinations, number of hospitals per 10 square kilometers, number of hospitals with PBSC collection per 10 square kilometers, number of hospitals with BM collection per 10 square kilometers, month dummies, and week dummies.

図3.5はサンプルを性別と年齢層（30歳未満か否か）で4つのサブセットに分割し、各サブセットでメッセージ効果を推定している。その結果、実験群Bは20代男性の確認検査到達比率を8ポイント程度高めており、統計的に有意である。したがって、実験群Bは20代男性の提供意向を高めるだけでなく、それを強固にしている可能性がある。また、実験群Bは20代男性の候補者選定・最終同意に統計的に有意な影響を与えていない。ただし、提供への効果は4ポイント程度であり、統計的に10%水準で有意である（表4.4）。先に述べたように、この効果は我々の介入効果だけでなく、幹細胞移植の需要を反映していると考えられる。若い男性ほど移植成績が良いならば、この世代の幹細胞移植の需要は他の性・年代よりも高いかもしれない。提供への効果はこの影響を含んでいるかもしれない。

# 4 Discussion and Conclusions

本研究はJMDPに登録した潜在ドナーの提供意向を高めるための情報提供の効果を検証した。その結果、一人の患者あたりのHLA適合ドナーが少ないという情報（Probability message）は20代男性の提供意向を10ポイント程度高め、確認検査への到達率も8ポイント程度高めている。人数に換算すると、この情報提供は提供に積極的な人を600人から756人に増やし、確認検査への到達した人を346人から462人に増やしている。さらに、幹細胞移植の需要の影響を含んでいるが、この情報はドナーの移植率を5ポイント高めている。人数に換算すると、Probability messageは71人から121人に増やしている。ただし、早いコーディネーションが患者の移植率を高めるという情報（Early Coordination message）を同時に提供すると、彼らはこのメッセージに比重を置いて意思決定するので、Probability messageの正の効果は消えてしまう。これらの結果は20代男性がHLA適合ドナーの数について過度な期待を持っており、フリーライド行動を取っていることを示唆している。

他の性・年代でProbability messageが統計的に有意な影響を与えていない原因は二つあると考えられる。第一に、20代男性と比較して、他の人はHLA適合ドナーの数を正しく見積もっているかもしれない。この場合、Probability messageは潜在ドナーの意思決定に全く影響を与えないはずである。第二に、20代男性と他の性・年代の利他的な選好が異なるという可能性である。経済学における利他行動の研究では、利他行動の動機が自身の利他行動そのもので効用を得るウォームグローと公共財の生産など利他行動の結果で効用を得る純粋利他性の二つに大別される(Andreoni 1990)。ウォームグローが相対的に強い人ほど、他者の行動に関心がないので、フリーライド行動を取る可能性が低くなる。したがって、20代男性と比較して、他の人はウォームグロー選好が利他行動の主動力となっているのかもしれない。要するに、Probability messageの異質性は信念の違いか動機の違いかで説明ができると考えられる。

また、Early Coordination messageは20代女性の全体の返信率に影響を与えていないが、短期間の返信（4日以内）に正の影響を与えている。これは返信行動そのものを促進しているよりむしろ、返信日数を短くしていることを示唆している。我々のデータ制約上、このメッセージが患者の移植までの期間を短くしているかどうかを検証できないが、この情報が少なくともコーディネーション期間の短縮に貢献している。他の性・年代で効果がなかったのは、他の人がこの情報をもともと持っている可能性が考えられる。

本研究はフィールド実験によって情報提供の因果効果をクリアに識別できるが、先に述べたメカニズムを識別できるようなデータ・実験設計ではないことに限界がある。しかしながら、本研究は実務的な含蓄がある。第1節で述べたように、造血幹細胞移植は提供の意向を示す時点と実際の提供の時点が異なるので、潜在ドナーの意向を維持することが骨髄ドナープログラムの課題となっている。特に、移植成績の良い若年ドナーが意向を維持できず、脱落しやすい。本研究は患者とマッチしたときに、骨髄ドナープログラムの事務局が与える情報によって若年潜在ドナーの行動変容を促せる可能性を提示しており、この課題を解決するための施策の一つとなるだろう。とくに、他者のHLA適合情報は、移植成績の良いドナーがコーディネーションに進みやすくなるという意味で、コーディネーションの効率性を高めるだろう。しかしながら、情報提供が効果をないことを示す研究(e.g. Galen E. Switzer et al. 2018)もあるので、他国の骨髄ドナープログラムでの情報提供の効果検証が期待される。

# Appendix

Table 4.1: Logit Model of Reply and Intention

|  | | | Intention | | | |
| --- | --- | --- | --- | --- | --- | --- |
|  | Reply | | Positive | | Negative | |
|  | (1) | (2) | (3) | (4) | (5) | (6) |
| Treatment B | 1.112 | 1.149 | 1.093 | 1.092 | 0.950 | 0.969 |
|  | [0.940, 1.315] | [0.965, 1.367] | [0.982, 1.216] | [0.977, 1.220] | [0.848, 1.063] | [0.863, 1.089] |
| Treatment C | 0.947 | 1.023 | 0.979 | 0.988 | 0.998 | 1.026 |
|  | [0.801, 1.120] | [0.856, 1.222] | [0.878, 1.092] | [0.879, 1.109] | [0.889, 1.120] | [0.909, 1.160] |
| Treatment D | 1.060 | 1.061 | 1.019 | 1.024 | 1.005 | 1.002 |
|  | [0.893, 1.257] | [0.887, 1.268] | [0.914, 1.136] | [0.914, 1.148] | [0.896, 1.128] | [0.889, 1.129] |
| Covariates |  | X |  | X |  | X |
| Num.Obs. | 10985 | 10985 | 10985 | 10985 | 10985 | 10985 |
| Log.Lik. | -3884.517 | -3712.289 | -7534.803 | -7364.638 | -6945.023 | -6869.968 |
| Notes: We show odds ratios and associated 95 percent confidential intervals. Covariates are gender, squared polynomial of (demeaned) age, number of past coordinations, number of hospitals per 10 square kilometers, number of hospitals with PBSC collection per 10 square kilometers, number of hospitals with BM collection per 10 square kilometers, month dummies, and week dummies. | | | | | | |

Table 4.2: Logit Model of Coordination

|  | CT | | Candidate | | Consent | | Donation | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Treatment B | 1.189 | 1.197 | 1.071 | 1.042 | 1.093 | 1.059 | 1.074 | 1.036 |
|  | [1.049, 1.348] | [1.050, 1.363] | [0.860, 1.334] | [0.829, 1.310] | [0.867, 1.378] | [0.831, 1.348] | [0.833, 1.385] | [0.795, 1.351] |
| Treatment C | 1.083 | 1.081 | 1.014 | 0.967 | 1.038 | 0.988 | 1.030 | 0.976 |
|  | [0.951, 1.233] | [0.942, 1.242] | [0.808, 1.272] | [0.760, 1.229] | [0.816, 1.319] | [0.766, 1.274] | [0.792, 1.339] | [0.739, 1.288] |
| Treatment D | 1.150 | 1.179 | 1.119 | 1.123 | 1.157 | 1.156 | 1.055 | 1.053 |
|  | [1.011, 1.308] | [1.031, 1.349] | [0.895, 1.397] | [0.892, 1.415] | [0.915, 1.463] | [0.907, 1.475] | [0.812, 1.371] | [0.804, 1.380] |
| Covariates |  | X |  | X |  | X |  | X |
| Num.Obs. | 10435 | 10435 | 8587 | 8587 | 8558 | 8558 | 8441 | 8441 |
| Log.Lik. | -5909.753 | -5764.480 | -2427.295 | -2349.439 | -2243.901 | -2168.120 | -1906.131 | -1851.371 |
| Notes: We show odds ratios and associated 95 percent confidential intervals. Covariates are gender, squared polynomial of (demeaned) age, number of past coordinations, number of hospitals per 10 square kilometers, number of hospitals with PBSC collection per 10 square kilometers, number of hospitals with BM collection per 10 square kilometers, month dummies, and week dummies. | | | | | | | | |

Table 4.3: Linear Probability Model of Reply and Intentions for Males Less Then 30 Years

|  | | | Intention | | | |
| --- | --- | --- | --- | --- | --- | --- |
|  | Reply | | Positive | | Negative | |
|  | (1) | (2) | (3) | (4) | (5) | (6) |
| Treatment B | 0.0661\*\* | 0.0671\*\*\* | 0.0935\*\* | 0.0954\*\*\* | -0.0274 | -0.0283\*\* |
|  | (0.0237) | (0.0225) | (0.0346) | (0.0269) | (0.0182) | (0.0133) |
| Treatment C | 0.0197 | 0.0266 | 0.0218 | 0.0260 | -0.0021 | 0.0006 |
|  | (0.0258) | (0.0217) | (0.0314) | (0.0189) | (0.0230) | (0.0139) |
| Treatment D | 0.0620\*\* | 0.0658\*\* | 0.0198 | 0.0289 | 0.0422 | 0.0370\*\* |
|  | (0.0260) | (0.0258) | (0.0435) | (0.0253) | (0.0299) | (0.0176) |
| Control average | 0.7444 | 0.7444 | 0.3861 | 0.3861 | 0.3583 | 0.3583 |
| Covariates |  | X |  | X |  | X |
| Num.Obs. | 1556 | 1556 | 1556 | 1556 | 1556 | 1556 |
| Notes: \* p < 0.1, \*\* p < 0.05, \*\*\* p < 0.01. The robust standard errors are reported in parentheses. Covariates are gender, squared polynomial of (demeaned) age, number of past coordinations, number of hospitals per 10 square kilometers, number of hospitals with PBSC collection per 10 square kilometers, number of hospitals with BM collection per 10 square kilometers, month dummies, and week dummies. | | | | | | |

Table 4.4: Linear Probability Model of Coordination among Males

|  | CT | | Candidate | | Consent | | Donation | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Treatment B | 0.0589\* | 0.0769\*\* | 0.0327\* | 0.0380\*\* | 0.0283\* | 0.0330\* | 0.0388\*\* | 0.0426\*\* |
|  | (0.0315) | (0.0277) | (0.0189) | (0.0170) | (0.0158) | (0.0164) | (0.0146) | (0.0164) |
| Treatment C | -0.0196 | -0.0027 | -0.0174 | -0.0104 | -0.0074 | -0.0037 | 0.0034 | 0.0061 |
|  | (0.0340) | (0.0197) | (0.0157) | (0.0111) | (0.0140) | (0.0114) | (0.0150) | (0.0118) |
| Treatment D | -0.0214 | -0.0205 | 0.0092 | 0.0126 | 0.0192 | 0.0211 | 0.0177 | 0.0215\* |
|  | (0.0344) | (0.0192) | (0.0181) | (0.0135) | (0.0187) | (0.0137) | (0.0160) | (0.0124) |
| Control average | 0.2386 | 0.2386 | 0.0927 | 0.0927 | 0.0767 | 0.0767 | 0.0574 | 0.0574 |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
| Covariates |  | X |  | X |  | X |  | X |
| Num.Obs. | 1499 | 1499 | 1275 | 1275 | 1269 | 1269 | 1253 | 1253 |
| Notes: \* p < 0.1, \*\* p < 0.05, \*\*\* p < 0.01. The robust standard errors are reported in parentheses. Covariates are gender, squared polynomial of (demeaned) age, number of past coordinations, number of hospitals per 10 square kilometers, number of hospitals with PBSC collection per 10 square kilometers, number of hospitals with BM collection per 10 square kilometers, month dummies, and week dummies. | | | | | | | | |

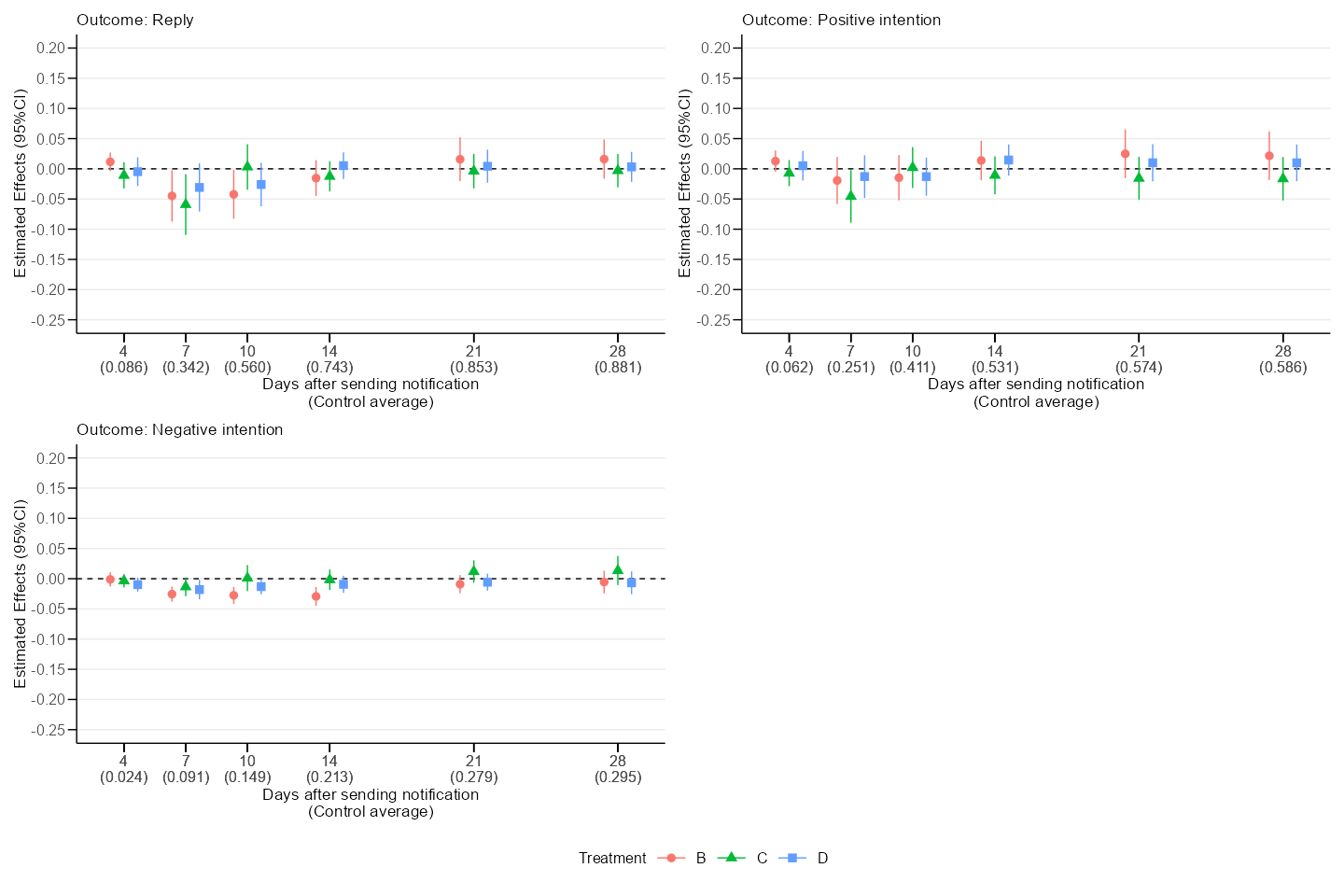


Figure 4.1: Effect on Reply within Specific Days after Sending Notification among Males More than 30. Notes: These plots show the average effect (and associated 95% confidential interval) on each outcome. We use robust standard errors We control number of past coordinations, number of hospitals per 10 square kilometers, number of hospitals with PBSC collection per 10 square kilometers, number of hospitals with BM collection per 10 square kilometers, month dummies, and week dummies.

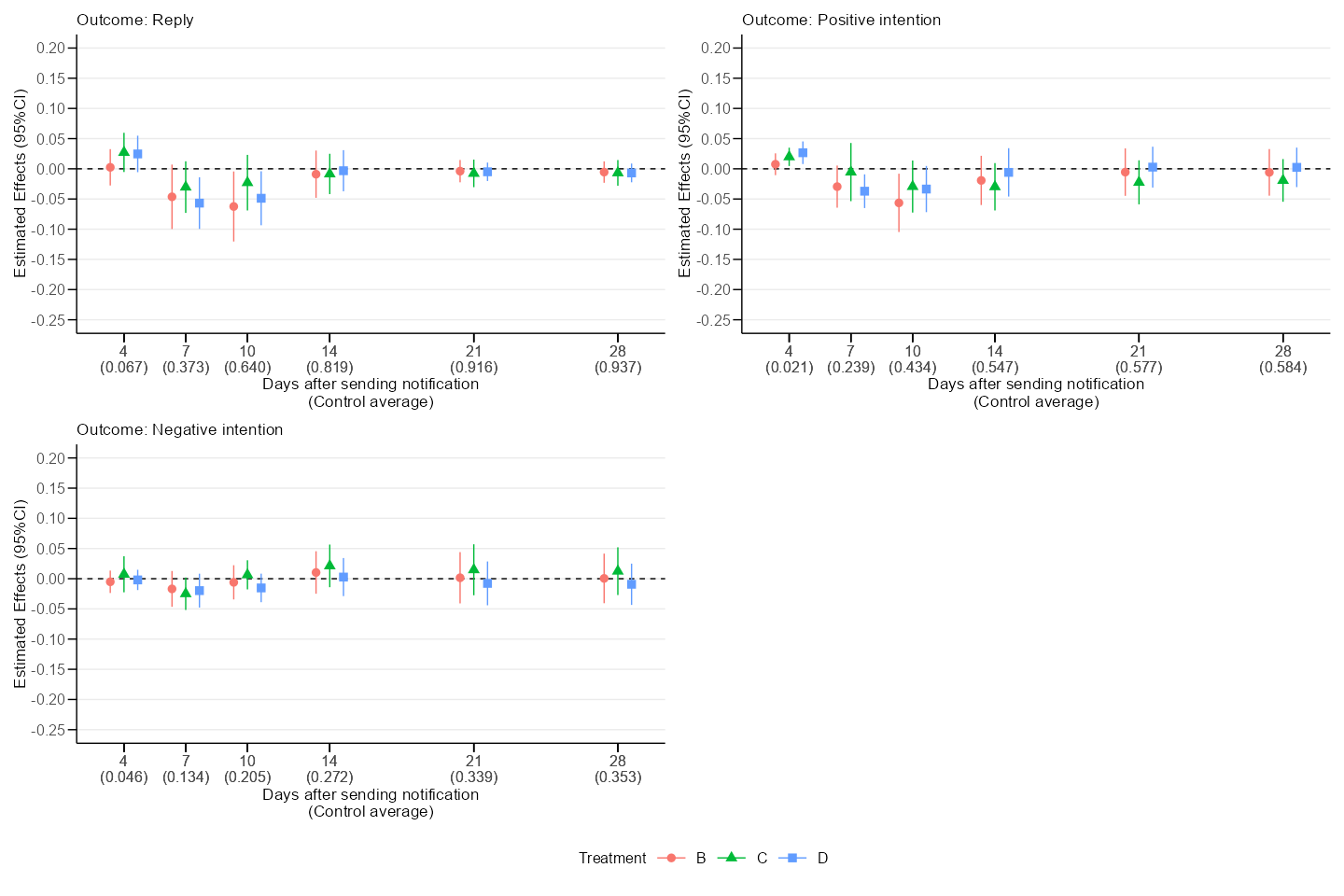


Figure 4.2: Effect on Reply within Specific Days after Sending Notification among Females More than 30. Notes: These plots show the average effect (and associated 95% confidential interval) on each outcome. We use robust standard errors We control number of past coordinations, number of hospitals per 10 square kilometers, number of hospitals with PBSC collection per 10 square kilometers, number of hospitals with BM collection per 10 square kilometers, month dummies, and week dummies.

Table 4.5: Conditional Average Treatment Effect Estimated by RCF

|  | Females | | | Males | | |
| --- | --- | --- | --- | --- | --- | --- |
|  | Estimate | S.E. | P-value | Estimate | S.E. | P-value |
| Age: Less than 30 | | | | | | |
| B | -0.007 | 0.048 | 0.884 | 0.123 | 0.040 | 0.002 |
| C | 0.042 | 0.048 | 0.379 | 0.048 | 0.041 | 0.245 |
| D | -0.058 | 0.049 | 0.236 | 0.044 | 0.040 | 0.278 |
| Age: More than or equal to 30 | | | | | | |
| B | -0.010 | 0.029 | 0.718 | 0.018 | 0.022 | 0.407 |
| C | -0.015 | 0.030 | 0.612 | -0.015 | 0.023 | 0.519 |
| D | 0.010 | 0.029 | 0.735 | 0.021 | 0.022 | 0.338 |
| Notes: See Athey and Wager (2019) for estimation method of conditional average treatment effect (CATE). Since these estimates are asymptotically normal, we calculate z-score under the null hypothesis that CATE is zero, and obtain p-value. | | | | | | |

# References

Andreoni, James. 1990. “Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving.” *The Economic Journal* 100 (401): 464. <https://doi.org/10.2307/2234133>.

Athey, Susan, and Guido Imbens. 2016. “Recursive Partitioning for Heterogeneous Causal Effects.” *Proceedings of the National Academy of Sciences* 113 (27): 7353–60. <https://doi.org/10.1073/pnas.1510489113>.

Athey, Susan, Julie Tibshirani, and Stefan Wager. 2019. “Generalized Random Forests.” *The Annals of Statistics* 47 (2). <https://doi.org/10.1214/18-AOS1709>.

Athey, Susan, and Stefan Wager. 2019. “Estimating Treatment Effects with Causal Forests: An Application.” *Observational Studies* 5 (2): 37–51. <https://doi.org/10.1353/obs.2019.0001>.

Bergstrom, Theodore C, Rodney J Garratt, and Damien Sheehan-Connor. 2009. “One Chance in a Million: Altruism and the Bone Marrow Registry.” *American Economic Review* 99 (4): 1309–34. <https://doi.org/10.1257/aer.99.4.1309>.

Carlana, Michela, Eliana La Ferrara, and Paolo Pinotti. 2022. “Goals and Gaps: Educational Careers of Immigrant Children.” *Econometrica* 90 (1): 1–29. <https://doi.org/10.3982/ECTA17458>.

Davis, Jonathan M. V., and Sara B. Heller. 2017. “Using Causal Forests to Predict Treatment Heterogeneity: An Application to Summer Jobs.” *American Economic Review* 107 (5): 546–50. <https://doi.org/10.1257/aer.p20171000>.

Diekmann, Andreas. 1985. “Volunteer’s Dilemma.” *Journal of Conflict Resolution* 29 (4): 605–10. <https://doi.org/10.1177/0022002785029004003>.

———. 1986. “Volunteer’s Dilemma. A Social Trap Without a Dominant Strategy and Some Empirical Results.” In *Paradoxical Effects of Social Behavior*, edited by Andreas Diekmann and Peter Mitter, 187–97. Heidelberg: Physica Heidelberg. <https://doi.org/10.1007/978-3-642-95874-8_13>.

Franzen, Axel. 1999. “The Volunteer’s Dilemma: Theoretical Models and Empirical Evidence.” In *Resolving Social Dilemmas: Dynamic, Structural, and Intergroup Aspects*, edited by Margaret Foddy, Michael Smithson, Sherry Schneider, and Michael A. Hogg, First. New York: Psychology Press. <https://doi.org/10.4324/9780203776001>.

Haylock, Michael, Patrick Kampkötter, Mario Macis, Jürgen Sauter, Susanne Seitz, Robert Slonim, Daniel Wiesen, and Alexander Schmidt. 2022. “Improving the Availability of Unrelated Stem Cell Donors: Evidence from a Major Donor Registry.” 29857. Cambridge, MA: National Bureau of Economic Research. <https://doi.org/10.3386/w29857>.

Hirakawa, Tsuneaki, Saiko Kurosawa, Kinuko Tajima, Yusuke Yamazaki, Nami Ikeda, Hiroto Kojima, Hidenori Tanaka, et al. 2018. “Japan Marrow Donor Program and its coordinating process: current situations.” *The Japanese Journal of Clinical Hematology* 59 (2): 153–60.

Kurosawa, Saiko, Yoshimi Tomine, Kazuhiro Yoshiuchi, Takahiro Fukuda, and Japan Marrow Donor Program. 2022. “An interview survey of psychosocial factors in unrelated stem cell donor candidates.” *Japanese Journal of Transplantation and Cellular Therapy* 11 (3): 169–76. <https://doi.org/10.7889/tct-21-024>.

Lacetera, Nicola, and Mario Macis. 2010. “Social Image Concerns and Prosocial Behavior: Field Evidence from a Nonlinear Incentive Scheme.” *Journal of Economic Behavior & Organization* 76 (2): 225–37. <https://doi.org/10.1016/j.jebo.2010.08.007>.

Lacetera, Nicola, Mario Macis, and Sarah S. Stith. 2014. “Removing Financial Barriers to Organ and Bone Marrow Donation: The Effect of Leave and Tax Legislation in the U.S.” *Journal of Health Economics* 33 (January): 43–56. <https://doi.org/10.1016/j.jhealeco.2013.10.006>.

Mellström, Carl, and Magnus Johannesson. 2008. “Crowding Out in Blood Donation: Was Titmuss Right?” *Journal of the European Economic Association* 6 (4): 845–63. <https://doi.org/10.1162/JEEA.2008.6.4.845>.

Murakami, Kayo, Hideki Shimada, Yoshiaki Ushifusa, and Takanori Ida. 2022. “Heterogeneous Treatment Effects of Nudge and Rebate: Causal Machine Learning Field Experiment on Electricity Conservation.” *International Economic Review*, June, iere.12589. <https://doi.org/10.1111/iere.12589>.

O’Donoghue, Ted, and M. Rabin. 2001. “Choice and Procrastination.” *The Quarterly Journal of Economics* 116 (1): 121–60. <https://doi.org/10.1162/003355301556365>.

Switzer, G. E., M. A. Dew, A. A. Stukas, J. M. Goycoolea, J Hegland, and R. G. Simmons. 1999. “Factors Associated with Attrition from a National Bone Marrow Registry.” *Bone Marrow Transplantation* 24: 313–19.

Switzer, Galen E., Mary Amanda Dew, Jean M. Goycoolea, Larissa Myaskovsky, Linda Abress, and Dennis L. Confer. 2004. “Attrition of Potential Bone Marrow Donors at Two Key Decision Points Leading to Donation.” *Transplantation* 77 (10): 1529–34. <https://doi.org/10.1097/01.TP.0000122219.35928.D6>.

Switzer, Galen E., Mario Macis, Rachel Fabi, Linda Abress, Dennis Confer, Jessica Bruce, Katie Howe, et al. 2018. “Providing Level-of-Match Information to Perfectly Matched Unrelated Stem Cell Donors: Evaluating Acceptability and Potential Changes in Donor Availability.” *Biology of Blood and Marrow Transplantation* 24 (10): 2110–18. <https://doi.org/10.1016/j.bbmt.2018.03.017>.

Takanashi, Minoko. 2016. “Marrow donor registration and cord blood banking: current issues.” *The Japanese Journal of Clinical Hematology* 57 (3): 271–77. <https://doi.org/10.11406/rinketsu.57.271>.

Wager, Stefan, and Susan Athey. 2018. “Estimation and Inference of Heterogeneous Treatment Effects Using Random Forests.” *Journal of the American Statistical Association* 113 (523): 1228–42. <https://doi.org/10.1080/01621459.2017.1319839>.

Wildman, John, and Bruce Hollingsworth. 2009. “Blood Donation and the Nature of Altruism.” *Journal of Health Economics* 28 (2): 492–503. <https://doi.org/10.1016/j.jhealeco.2008.11.005>.

1. 近年、ハプロ移植と呼ばれる、半合致したHLAを持つ近親者間での移植が広がりつつある。加えて、母子をつなぐさい帯や胎盤に含まれる血液細胞の移植（さい帯血移植）も同様に人気になっている。骨髄移植と異なり、HLAが完全に合致していなくても、さい帯血移植は実施できる。 [↑](#footnote-ref-20)
2. これは珍しいHLAの型を持つ新しいドナーの確率が低いからである。 [↑](#footnote-ref-21)
3. 20代男性のコーディネーションの15%がドナーの不健康（既往歴・腰痛・治療中など）で中断している一方で、41%が不健康以外のドナー理由（連絡が取れない・都合がつかないなど）で中断している。移植はドナーの家族の同意を必要とする。不健康以外のドナー理由は家族の同意を得られていないことを含んでいない。20代男性のコーディネーションの6%がこの理由で中断している。 [↑](#footnote-ref-22)
4. JMDPは返信目安が10日以内であるということを適合通知に明記している。 [↑](#footnote-ref-23)
5. JMDPは同時に潜在ドナーにJMDPが適合通知を送付したことを知らせるSNSメッセージを送付する。 [↑](#footnote-ref-25)
6. 介入メッセージを作成したとき、我々は潜在ドナーに対して過度な心理的プレッシャーを与えないように配慮している。具体的には、第一に、嘆願調のような表現を避けている。第二に、JMDPが公開している情報のみを用いている。また、移植リスクは通常通り説明している。 [↑](#footnote-ref-31)
7. 一人が海外に在住していた。また、データ提供時点でコーディネーションが進行中であった潜在ドナーは104名いた。コーディネーションが進行中の潜在ドナーの比率は実験群間でバランスしている（F-test, p-value = ）。 [↑](#footnote-ref-34)
8. <https://www.jmdp.or.jp/hospitals/view2/> (access date: August 4, 2022) [↑](#footnote-ref-35)
9. クラスター標準誤差を用いた回帰分析を頑健性の確認として実施し、本論で示す主要な結果に変化がないことを確認している。 [↑](#footnote-ref-38)
10. 20代女性について、%。30歳以上の男性について、%。30歳以上の女性について、%。 [↑](#footnote-ref-46)
11. 技術的な詳細は Athey and Imbens (2016) ・ Wager and Athey (2018) を参照せよ。また、本研究ではRCFを特殊ケースとして扱う一般化ランダムフォレストと呼ばれる手法を用いているが、直観は変わらない。一般化ランダムフォレストの詳細は Athey and Wager (2019) ・ Athey, Tibshirani, and Wager (2019) を参照せよ。 [↑](#footnote-ref-48)
12. 30歳以上の男女については補論の図4.1と図4.2に示す。これらのグループでは特徴的な結果がない。 [↑](#footnote-ref-59)