**Famílias excluídas do crédito em um país em desenvolvimento[[1]](#footnote-1)**

1. **Os objetivos da atividade**

* identificar famílias com restrições de crédito e excluídas do crédito
* criar variáveis fictícias (indicador)
* explicar por que o viés de seleção é uma questão importante

1. **Contextualização**

Os empréstimos, através de instituições formais ou informais, podem ajudar o consumo e possibilitar investimentos. No entanto, os credores não podem observar o quanto o mutuário está se esforçando para reembolsar o empréstimo e não podem condicionar o contrato de empréstimo a esse esforço. O relacionamento entre o credor e o devedor, como o empregador e o empregador, é chamado de problema de agente principal[[2]](#footnote-2).

Os credores podem mitigar ou eliminar parcialmente o risco de inadimplência, exigindo garantias, que podem ser recuperadas e vendidas para pagar o empréstimo, se o mutuário deixar de pagar. As pessoas que são incapazes de fornecer essa garantia geralmente precisam tomar empréstimos sob condições mais desfavoráveis ​​(taxas de juros mais altas) ou podem receber um empréstimo totalmente recusado. Chamamos o primeiro grupo de restrições de crédito e o último grupo de excluídos.

Como a capacidade de contrair empréstimos depende da riqueza e da renda de uma pessoa, essas restrições e exclusões de crédito contribuem para a desigualdade, e algumas oportunidades de crescimento econômico não são realizadas. Por exemplo, uma pessoa trabalhadora e sem ativos pode ter um empréstimo recusado para iniciar um negócio, o que poderia contribuir tanto para aumentar sua renda quanto para a atividade econômica.

Para projetar políticas que ajudem os mercados de crédito a funcionarem melhor, primeiro precisamos observar como as condições de financiamento e as fontes de financiamento disponíveis variam de acordo com as características da família. Às vezes, os mutuários excluídos dos mercados de crédito formais ainda podem obter empréstimos através de outros credores, como parentes ou amigos, por isso é importante considerar essas fontes ao classificar as pessoas como restritas de crédito ou excluídas.

Em países onde os mercados de crédito formais ainda estão em desenvolvimento, os acordos informais são formas importantes para as comunidades compartilharem recursos e agruparem os riscos. Por exemplo, na Etiópia, as famílias podem fazer parte de uma rede de apoio social denominada "*iddir*", um grupo de pessoas que pagam regularmente dinheiro para uma reserva comum que é partilhada entre os membros do grupo que dela necessitam.

Analisaremos os dados de uma pesquisa domiciliar etíope (o Levantamento Socioeconômico da Etiópia – the Ethiopian Socioeconomic Survey) para investigar as condições de financiamento que diferentes tipos de famílias enfrentam. Para além das famílias com restrições de crédito e excluídas do crédito, iremos também analisar um terceiro grupo de agregados familiares que são "mutuários desencorajados", o que significa que não solicitaram um empréstimo porque pensaram que seriam recusados.

1. **Famílias que não obtiveram um empréstimo**

Os dados do Inquérito Socioeconómico da Etiópia (ESS) (Ethiopian Socioeconomic Survey) foram recolhidos em 2013–14 a partir de uma amostra nacionalmente representativa de agregados familiares. As famílias foram questionadas sobre tópicos como as condições de moradia, ativos e acesso ao crédito.

Faça o download dos dados do ESS e do questionário da pesquisa:

* Baixe os [dados do ESS](https://tinyco.re/6067633). Leia a guia do dicionário de dados (Data dictionary) e verifique se você sabe o que cada variável representa. Adiante discutiremos exatamente como algumas dessas variáveis foram construídas.
* Para a documentação, acesse o site de [download de dados](http://tinyco.re/9757350).
* No cabeçalho "Materiais Relacionados" (Related Materials), no meio da página, faça o download do arquivo em PDF chamado "Pesquisa Socioeconômica Etíope 2013-2014, Questionário das Famílias" (2013–2014 Ethiopian Socioeconomic Survey, Household Questionnaire).

1. Os dados já estão em um formato relativamente limpo, portanto, começaremos resumindo as informações na guia "Todas as residências" (All households), começando com as características da região e da família.
   1. Crie uma Tabela Dinâmica mostrando a proporção de famílias que viveram em cada região e tipo de área, com "região" como variável de linha e "rural" como a variável de coluna.
   2. Use a função IF do Excel para criar uma variável que seja igual a 1 se o chefe do domicílio (família) for do sexo feminino e 0 se for do sexo masculino. Variáveis ​​que são codificadas dessa maneira são conhecidas como variáveis ​​dummy (ou variáveis ​​indicadoras)[[3]](#footnote-3). Qual percentagem de chefes de familiares eram mulheres?
   3. Crie uma tabela apropriada para as variáveis ​​"hhsize", "gender", "age", "young\_children", "working\_age\_adults", "max\_education" e "number\_assets".
   4. Escreva um pequeno parágrafo descrevendo as informações em suas tabelas para I (c).

Agora que temos uma ideia de como são nossos dados, vamos passar a identificar famílias potencialmente excluídas do mercado de crédito ou com restrições de crédito. As primeiras são famílias que acham impossível tomar empréstimos e as últimas são famílias que só podem tomar empréstimos em condições desfavoráveis.

As variáveis em nosso conjunto de dados relacionadas a esse problema são "did\_not\_apply" e "loan\_rejected". Em breve, veremos também as respostas dadas nas variáveis "reason\_not\_apply1" e "reason\_not\_apply2".

1. Usando a guia "Todas as residências" (All households):
   1. Crie uma tabela de frequência com "did\_not\_apply" como variável de linha e "loan\_rejected" como variável de coluna. Inclua os espaços em branco como uma linha separada.
   2. Analisando essas duas variáveis, explique por que algumas observações devem ser excluídas e removê-las do conjunto de dados. Também remova todos os domicílios com informações ausentes para uma ou mais dessas variáveis. Das observações não excluídas, que porcentagem de domicílios solicitou um empréstimo nos últimos 12 meses? Desses domicílios, qual porcentagem foi bem-sucedido?
   3. Para as categorias resultantes na tabela de frequência, explique se os domicílios nessa categoria podem ser descritos como restritos por crédito, excluídos por crédito ou ambos.

Para criar categorias operacionais para usar ao longo desta atividade, rotularemos as famílias como:

* Bem-sucedidos: famílias que solicitaram um empréstimo e receberam o empréstimo
* Negado: famílias que se inscreveram, mas não receberam o empréstimo
* Não se aplica: famílias que não solicitaram um empréstimo.

Deve ter em conta que as famílias "recusadas" são provavelmente apenas um subconjunto das famílias excluídas por crédito, uma vez que haverá famílias que são excluídas do crédito e nem sequer solicitam um empréstimo. Poder-se-ia, por exemplo, argumentar que os agregados familiares que responderam "Garantia inadequada" (Inadequate Collateral) ou "Não conhecem qualquer credor" (Do Not Know Any Lender) também são susceptíveis de serem excluídos do crédito.

1. Utilizando o subconjunto de dados da questão II, alínea b):
   1. Crie uma variável chamada "HH\_status" com as categorias acima.
   2. Crie uma variável "desfavorecido" que leva o valor 1 se a família não solicitou um empréstimo, porque acreditava que não receberia um empréstimo (a resposta "Acredite seria recusada" (Believe Would Be Refused) em "reason\_not\_apply1" ou "reason\_not\_apply2"). Quantas famílias (e que porcentagem) são desencorajadas pelos mutuários? Nota: esta é uma definição bastante restrita de "desencorajado" (discouraged) e pode-se facilmente argumentar que outros critérios também devem ser considerados sob este rótulo.
   3. Crie uma variável 'credit\_constrained' que leva o valor 1 (ou sim) para os domicílios que forneceram um motivo para não aplicar diferente de 'NA', 'Other' ou 'Have Properate Farm' em uma das duas perguntas 'reason\_not\_apply1' ou 'reason\_not\_apply2' e 0 caso contrário. Por exemplo, uma família que responda " Have Adequate Farm" em "reason\_not\_apply1" e " Do Not Know Any Lender" não será classificada como restrição de crédito. Quantas residências (e que porcentagem) têm restrições de crédito?
   4. Crie uma tabela de frequências mostrando a razão mais importante para não solicitar um empréstimo e outra mostrando a segunda razão mais importante para não se candidatar. Quais foram os motivos mais comuns para não se inscrever?

Analisaremos agora as razões declaradas para querer um empréstimo, comparando as famílias que obtiveram sucesso ("HH\_status" igual a "success") com aquelas que não obtiveram sucesso ("HH\_status" igual a "denied").

1. Para ambos os grupos, crie uma tabela mostrando a proporção de famílias para cada finalidade do empréstimo. Você perceberá que na guia ' All households' o motivo de todos os empréstimos bem-sucedidos é 'Other'. Por esse motivo, você deve usar a guia ' Got loan' para recuperar os motivos das informações do empréstimo para empréstimos bem-sucedidos.) o propósito de empréstimos para mutuários negados e bem-sucedidos semelhantes? Dica: pode ajudar pensar nas categorias amplas de gastos com consumo e investimento.
2. Usando as informações na guia " All households" e na guia " Got loan", para cada grupo de domicílios:
   1. Crie uma tabela, como mostrado na Figura 1, para comparar as médias das características domésticas especificadas.

Figura 1 – Características dos mutuários excluídos e bem-sucedidos

| **Household characteristic** | **Successful borrowers** | **Credit-excluded** |
| --- | --- | --- |
| Age of household head |  |  |
| Highest education in household |  |  |
| Number of assets |  |  |
| Household size |  |  |
| Number of young children |  |  |
| Number of working-age adults |  |  |

* 1. Para cada característica, explique como isso pode afetar a capacidade de uma família obter um empréstimo (ceteris paribus).
  2. Discuta se você vê esse padrão nos dados.
  3. Agora tente condicionar a variável "rural" ou "região" e discuta como (se houver) seus resultados mudam.

1. Usando a Figura 1, sem condicionamento em "rural" ou "região", execute o seguinte:
   1. Calcule a diferença nas médias (mutuários "bem-sucedidos" menos os mutuários "recusados").
   2. Estabeleça se as diferenças entre os dois grupos são estatisticamente significativas no nível de confiança de 5%, calculando o intervalo de confiança de 95% para a diferença de médias entre os dois subgrupos (‘êxito’ menos ‘negado’). Dica: use a função CONFIDENCE.T do Excel.
   3. Plote um gráfico de colunas mostrando as diferenças no eixo vertical (classificadas do menor para o maior) e as características das famílias no eixo horizontal. Adicione os intervalos de confiança da pergunta VI (b) ao gráfico.
   4. Interprete suas descobertas.
2. Usando as informações na guia " All households":
   1. Crie uma tabela semelhante à Figura 1, mas com colunas adicionais para mutuários desanimados e famílias com restrições de crédito.
   2. Compare os meios entre os quatro grupos (você não precisa avaliar a significância estatística formalmente).

Um [estudo sobre o acesso a empréstimos na Etiópia](https://tinyco.re/9835958) analisou a relação entre o montante do empréstimo e as características do agregado familiar. Ao fazer isso, eles precisavam explicar o viés de seleção[[4]](#footnote-4), porque só observamos montantes positivos de empréstimos para mutuários bem-sucedidos. Se tivéssemos apenas dados para os mutuários bem-sucedidos, então a nossa amostra não seria representativa da população de interesse (todas as famílias), por isso teríamos que interpretar os nossos resultados com cautela. No nosso caso, temos informações sobre todos os domicílios, para que possamos comparar as características observáveis ​​para ver se os mutuários bem-sucedidos são semelhantes aos outros domicílios.

1. Pense em outro exemplo em que pode haver viés de seleção, em outras palavras, em que os dados que observamos não são representativos da população de interesse.
2. **Famílias que receberam um empréstimo**

Para as famílias que obtiveram um empréstimo com sucesso, veremos:

* finalidade do empréstimo
* duração do(s) empréstimo(s)
* montante do empréstimo e taxa de juros cobrada
* de quem a família pediu emprestado.

Também veremos se há algum relacionamento entre essas características do empréstimo e as características da família.

Agora, usaremos as variáveis relacionadas às datas de início e término do empréstimo para calcular a duração do empréstimo. Antes de usar essas variáveis, precisamos verificar se as entradas das variáveis fazem sentido. Algumas dessas informações podem ser gravadas incorretamente (por exemplo, o ano está faltando um dígito ou o mês é um número e não uma palavra).

1. Usando a guia "Got loan", faça o seguinte:
   1. Verifique as variáveis 'loan\_startmonth', 'loan\_startyear', 'loan\_endmonth' e 'loan\_endyear' e substitua as entradas gravadas incorretamente pela entrada correta (se possível) ou em branco (se não for possível inferir a entrada correta). Observação: algumas entradas são registradas como 'Pagume', que corresponde ao início de setembro no calendário etíope.
   2. Para calcular a duração do empréstimo, precisamos combinar as variáveis de mês e ano em uma variável de data. Use a função CONCATENATE do Excel para criar variáveis para a data de início e término e formatá-las como variáveis de data.
   3. Faltam algumas datas (meses ou anos). Calcule a porcentagem dos dados que estão faltando e explique se você acredita que os dados ausentes são um problema sério.
   4. Crie uma variável contendo a duração do empréstimo (data de término menos data de início), que será medida em dias.
   5. Você notará que algumas datas foram gravadas incorretamente, com a data de início posterior à data de término. Poderíamos tratar essas entradas como ausentes ou trocar as datas de início e término. Crie duas variáveis para a duração do empréstimo, uma com todas as entradas negativas registradas em branco e outra com entradas negativas substituídas como números positivos. Dica: a função ABS do Excel converte qualquer número em um número positivo.
   6. Para esta atividade, definiremos um empréstimo de longo prazo como durando mais de um ano (365 dias), que usaremos em perguntas posteriores. Para esta definição, use a variável duração do empréstimo que converte valores negativos para positivos. Crie uma variável de indicador chamada "long\_term" que seja igual a 1 se o empréstimo for de longo prazo e 0 caso contrário. Qual a porcentagem de empréstimos em longo prazo?
   7. **Criando e formatando variáveis de tempo**
2. Usando as variáveis ​​'loan\_amount' e 'loan\_interest':
   1. Crie tabelas para apresentar a distribuição do valor do empréstimo (média, desvio padrão, máximo e mínimo), um usando o valor do empréstimo (loan\_amount) e outro usando o valor total a ser reembolsado (loan\_amount + interest). Exclua a observação anteriormente identificada como tendo uma taxa de juros extremamente alta. Lembre-se de atribuir títulos às suas tabelas. Descreva todos os recursos dos dados que você achar interessantes.
   2. Como mencionado anteriormente, a taxa de juros é uma condição de empréstimo que pode variar amplamente entre as famílias. Aqui, consideraremos a taxa de juros como sendo os juros pagos como uma porcentagem do valor do empréstimo. Calcule a taxa de juros para cada empréstimo nos dados. Exclua as observações em que os juros pagos não sejam registrados.
   3. Verifique valores extremos (taxas de juros muito grandes ou zero). Você também pode criar um gráfico de dispersão (com taxa de juros no eixo vertical e valor do empréstimo no eixo horizontal) para ajudar a identificar observações extremas. Exclua a observação com a taxa de juros mais extrema de outros cálculos. Qual a porcentagem dos empréstimos sem juros?
   4. Faça tabelas para apresentar as médias, máximas, mínimas e quartis do valor do empréstimo e da taxa de juros, calculando essas medidas separadamente para empréstimos de longo e curto prazo. As distribuições de taxas de juros para empréstimos de curto e longo prazo parecem diferentes?
   5. Crie uma tabela mostrando a correlação entre a taxa de juros e as características da família. Interpretando a taxa de juros cobrada como uma medida de risco de inadimplência (incapacidade de pagar), explique se as relações implícitas nos coeficientes são as esperadas (por exemplo, você espera que as taxas de juros sejam mais altas para famílias com menos ativos, mais dependentes etc.).
3. Agora, examinaremos as fontes de financiamento e como elas estão relacionadas às características do empréstimo.
   1. Crie uma tabela mostrando a proporção de empréstimos (em termos da variável da coluna) com a fonte de financiamento ("borrowed\_from") como a variável de linha e "rural" como a variável da coluna. Crie uma tabela semelhante, mas com "borrowed\_from\_other" como a variável de linha. Parece que as famílias rurais usam diferentes fontes de financiamento das famílias urbanas? Dica: pode ajudar a pensar em fontes de financiamento em termos de instituições formais, informais e outras, como microfinanciadores ou ONGs.)
   2. Para cada uma das variáveis abaixo, crie uma tabela mostrando a média dessa variável, com 'borrowed\_from' como variável de linha e 'rural' como variável de coluna. Comente as semelhanças ou diferenças entre linhas e colunas que achar interessantes e sugira explicações para o que você observa.

* duração do empréstimo (usando a variável na qual durações negativas foram transformadas em durações positivas)
* montante do empréstimo
* taxa de juro
  1. Crie uma tabela que mostre a proporção de 'gender' (em termos da variável de linha) com 'borrowed\_from' como variável de linha e 'rural' como variável de coluna. Parece haver uma relação entre o gênero do chefe de família e os tipos de financiamento utilizados?
  2. Quais outras variáveis atualmente não estão em nosso conjunto de dados, mas também podem ser importantes para nossa análise nas perguntas II e III?

1. Nesta atividade, examinamos padrões de empréstimos e acesso ao crédito, mas não somos capazes de fazer declarações causais como 'mudanças em X causam restrição de crédito às famílias' ou 'característica Y causa melhor acesso ao crédito'. Descreva uma política que possa ajudar a melhorar o acesso das famílias a empréstimos e como projetar a implementação para que você possa avaliar os efeitos causais dessa política.

1. Preparada com base em Doing Economics: empirical projects (http://www.core-econ.org/doing-economics/index.html). [↑](#footnote-ref-1)
2. É um relacionamento assimétrico no qual uma parte (principal) se beneficia de alguma ação ou atributo da outra parte (o agente) sobre a qual as informações do principal não são suficientes para impor um contrato completo. [↑](#footnote-ref-2)
3. Variável dummy (variável do indicador) é uma variável que recebe o valor 1 se uma determinada condição for atendida e 0 caso contrário. [↑](#footnote-ref-3)
4. Um problema que ocorre quando a amostra ou os dados observados não são representativos da população de interesse. Por exemplo, indivíduos com certas características podem ter maior probabilidade de fazer parte da amostra observada (por exemplo, os alunos têm mais probabilidade do que os CEOs de participar de experimentos em laboratórios de informática). [↑](#footnote-ref-4)