[Attachment Removed] Fw: RE: Re[2]: FW: Re[2]: [EXTERNAL] Question about access to data from 2018 DI ASR

18 messages

Janet Stein <jbstein@nber.org> Thu, May 7, 2020 at 2:34 PM
Reply-To: Janet Stein <jbstein@nber.org>
To: David Molitor <dmolitor@illinois.edu>
Cc: Nicole Maestas <maestas@hcp.med.harvard.edu>, Dick Woodbury <dickwoodbury1@gmail.com>

Hi David,
I hope all is well. I just received this message from SSA with the data you had requested. (This may well be the quickest data turnaround of all time!)
Let me know if you have any questions.
Janet

------ Forwarded Message ------
From: "Jankowski, John" <John.Jankowski@ssa.gov>
To: "Janet Stein" <jbstein@nber.org>
Sent: 5/7/2020 2:15:57 PM
Subject: RE: Re[2]: FW: Re[2]: [EXTERNAL] Question about access to data from 2018 DI ASR

Hi Janet,

I just received the disclosure-reviewed tables that David requested. There are two Zip files attached: (1) the first one provides tables for certain age groupings (e.g., less than 20, 20-24, 25-29, etc.); (2) the second one provides tables for individual ages from 45 to 64. It should be clear from the file names what each table covers. All tables cover the period 2008-2017.

Hopefully this is everything they will need for their project. If they have any questions, just let me know.

Thanks,
John

-----Original Message-----
From: Janet Stein <jbstein@nber.org>
Sent: Monday, March 16, 2020 12:21 PM
To: Jankowski, John <John.Jankowski@ssa.gov>
Subject: Re[2]: FW: Re[2]: [EXTERNAL] Question about access to data from 2018 DI ASR

Hi John,
Ready by end-April? That counts as a miracle! I'll let David know.
Janet

------ Original Message ------
From: "Jankowski, John" <John.Jankowski@ssa.gov>
To: "Janet Stein" <jbstein@nber.org>
Sent: 3/16/2020 1:16:54 PM
Subject: RE: FW: Re[2]: [EXTERNAL] Question about access to data from 2018 DI ASR

> Hi Janet,
>
> I heard back from our Baltimore staff on this. They can have one of our contractors put together the data that David is requesting, but it likely won't be ready until end-April. Let me know if this timeframe will work for him.
>
> Thanks,
>
> John
>
> From: Janet Stein <jbstein@nber.org>
> Sent: Tuesday, March 10, 2020 12:05 PM
> To: Jankowski, John <John.Jankowski@ssa.gov>

Subject: Re: FW: Re[2]: [EXTERNAL] Question about access to data from 2018 DI ASR

Thanks, John

Here's what I heard back from David Molitor:

Thanks for the update. Limiting the years to 2008 – 2017 is fine. The essence of what we're doing is comparing how things change as you cross grid rule age thresholds, so having something more granular like year of age near these thresholds is critical, hence our request for year of age between 45 and 64.

Is it possible to get more clarity on the disclosure rule for suppressing cell values (e.g., if value is < 11)? It is possible that simply knowing that the value is less than this threshold will be sufficient for our needs.

In the event that cell suppression under year-of-age splits creates a serious problem for our analysis, one solution might be to aggregate data by calendar year rather than by age. A downside here is that this would prevent us from analyzing how grid rule patterns change over the business cycle, but at least we would be able to conduct our primary analysis.

------ Original Message ------

From: "Jankowski, John" <John.Jankowski@ssa.gov>

To: "Janet Stein" <jbstein@nber.org>

Sent: 3/10/2020 2:51:22 PM

Subject: FW: Re[2]: [EXTERNAL] Question about access to data from 2018 DI ASR

> Hi Janet,
>
> Just following up on this one. Let me know if you want me to pursue this request.
>
> Thanks,
>
> John

From: Jankowski, John
Sent: Wednesday, March 04, 2020 8:39 AM
To: 'Janet Stein' <jbstein@nber.org>
Subject: RE: Re[2]: [EXTERNAL] Question about access to data from 2018 DI ASR

Hi Janet,

I spoke to someone from our Baltimore office this morning about this request. We have contractors up there who should be able to put this together, though there are some limitations: (1) breaking the data into individual ages will lead to a lot of cells being suppressed (due to disclosure rules); and (2) because of the way the dataset is structured, it is only possible in the short-term to provide data for the last 10 years. (As I understand it, the dataset only maintains the past 10 years of data; for years 11 and later, each year is made into a separate dataset, meaning that they would have to merge the current file with many separate dataset. This would be very time consuming.)

As a result of these limitations, this is what Baltimore proposes pursuing:

“I plan to pursue the feasibility of providing customized versions of the 2018 Annual Statistical Report on the Social Security Disability Insurance Program Tables 60-63.

The customization would include

A. Limiting the years to 2008 – 2017

B. Using age groupings:

1. Less than 20

2. 20 – 24

3. 25 – 29

4. 30 – 34

5. 35 – 39

6. 40 – 44

7. 45 – 49

8. 50 – 54

9. 55 – 59

10. 60 – 64

11. 65 and over

There would be 4 tables for each age grouping for a total of 44 tables.

The tables would be subject to our disclosure review process."

Please run this by the PIs and let me know if this will work for them.

Thanks,

John

From: Janet Stein <jbstein@nber.org>
Sent: Friday, February 28, 2020 10:22 AM
To: Jankowski, John <John.Jankowski@ssa.gov>
Subject: Re[2]: [EXTERNAL] Question about access to data from 2018 DI ASR

Hi John,

Here's the reply:

We would like splits by year of age. If it's easier or important for SSA to bin some ages, the following bins would suffice:

Ages <20,

Ages 20-24,

Ages 25-29,

Ages 30-34,

Ages 35-39,

Ages 40-44,

year of age starting at 45 (i.e., 45, 46, ..., 64)

------ Original Message ------

From: "Jankowski, John" <John.Jankowski@ssa.gov>

To: "Janet Stein" <jbstein@nber.org>

Cc: "Murphy, John" <John.Murphy@ssa.gov>

Sent: 2/28/2020 11:26:56 AM

Subject: RE: [EXTERNAL] Question about access to data from 2018 DI ASR

> Hi Janet,
>
> I received a response on this data request. Could you ask the PIs for some clarification?
>
> “My understanding of what is written below is that the requestor would like 2018 Annual Statistical Report on the Social Security Disability Insurance Program Tables 60-63 by age. I need to know what age splits they want. Are they interested in the standard age splits?:

- Under age 18

- Aged 18–64

- Aged 65 or older”

Thanks,

John

From: Janet Stein <jbstein@nber.org>
Sent: Tuesday, February 25, 2020 8:53 AM
To: Murphy, John <John.Murphy@ssa.gov>; Jankowski, John <John.Jankowski@ssa.gov>
Subject: [EXTERNAL] Question about access to data from 2018 DI ASR

Hi John and John,

David Molitor is preparing his WIP at SSA on March 30th (Project NB20-01, Health Variations Among DI Recipients: Evidence from Medicare). Their team has asked about the possibility of accessing some aggregated data from Tables 60-63 of SSA's Annual Statistical Report before then.

The description of what they would like is below. Can you let me know if this request is feasible and/or how they should proceed?

Thanks,

Janet

Here is a description of what they would like to obtain:

Reference: Tables from the 2018 DI ASR "Outcomes of Applications for Disability Benefits"
https://www.ssa.gov/policy/docs/statcomps/di_asr/2018/sect04.pdf

Beneficiary categories: both "all disabled beneficiaries" and "disabled workers"

Years: all available (1992+ for all disabled beneficiaries, 1999+ for disabled workers)

Data elements requested:
Table 60, "Total"
Table 60, "Pending final decision"
Table 60, "Technical denials"
Table 60, "Medical decisions, Denials, Medical"
Table 60, "Medical decisions, Denials, Subsequent nonmedical"
Table 60, "Medical decisions, Allowances, Awards"
Table 60, "Medical decisions, Allowances, Subsequent denials"
Table 61, "All decisions, Total"
Table 61, "All decisions, Allowances"
Table 62, "All decisions, Total"
Table 62, "All decisions, Allowances"
Table 63, "All decisions, Total"
Table 63, "All decisions , Allowances"

Requesting records by Year and Age (at the time of application)

SSDI_ASR_CellSupp - Ages 45-64.zip
213K

David Molitor <dmolitor@illinois.edu> Thu, May 7, 2020 at 2:35 PM

To: Colleen Carey <colleen.carey@cornell.edu>, "Miller, Nolan H" <nmiller@illinois.edu>

Great timing!
[Quoted text hidden]
--

David Molitor
Assistant Professor of Finance and Economics
Gies College of Business, University of Illinois
www.davidmolitor.com

SSDI_ASR_CellSupp - Ages 45-64.zip
213K

David Molitor <dmolitor@illinois.edu> Thu, May 7, 2020 at 3:16 PM
To: Janet Stein <jbstein@nber.org>
Cc: Nicole Maestas <maestas@hcp.med.harvard.edu>, Dick Woodbury <dickwoodbury1@gmail.com>

Fantastic news! I let you know if I have any questions after we dig in.

Thanks so much,
David
[Quoted text hidden]

David Molitor <dmolitor@illinois.edu> Thu, May 7, 2020 at 3:38 PM
To: Janet Stein <jbstein@nber.org>

Hi Janet,

One quick thing. John Jankowski mentions that two ZIP files are attached, but only one came through for me (SSDI_ASR_CellSupp - Ages 45-64.zip). It sound like there should also be a separate file for the other age groupings (less than 20, 20-24, 25-29, etc.) . Is that available?

Thanks,
David

On Thu, May 7, 2020 at 2:34 PM Janet Stein <jbstein@nber.org> wrote:
[Quoted text hidden]

Janet Stein <jbstein@nber.org> Thu, May 7, 2020 at 3:41 PM
Reply-To: Janet Stein <jbstein@nber.org>
To: David Molitor <dmolitor@illinois.edu>

I sent you both. I'll try again separately.
[Quoted text hidden]

Janet Stein <jbstein@nber.org> Thu, May 7, 2020 at 3:41 PM
Reply-To: Janet Stein <jbstein@nber.org>
To: David Molitor <dmolitor@illinois.edu>

Second attempt with just the missing attachment.

------ Forwarded Message ------
From: "Jankowski, John" <John.Jankowski@ssa.gov>
To: "Janet Stein" <jbstein@nber.org>
Sent: 5/7/2020 2:15:57 PM
Subject: RE: Re[2]: FW: Re[2]: [EXTERNAL] Question about access to data from 2018 DI ASR

Hi Janet,

I just received the disclosure-reviewed tables that David requested. There are two Zip files attached: (1) the first one provides tables for certain age groupings (e.g., less than 20, 20-24, 25-29, etc.); (2) the second one provides tables for individual ages from 45 to 64. It should be clear from the file names what each table covers. All tables cover the period 2008-2017.

Hopefully this is everything they will need for their project. If they have any questions, just let me know.

Thanks,
John

-----Original Message-----
From: Janet Stein <jbstein@nber.org>
Sent: Monday, March 16, 2020 12:21 PM
To: Jankowski, John <John.Jankowski@ssa.gov>
Subject: Re[2]: FW: Re[2]: [EXTERNAL] Question about access to data from 2018 DI ASR

Hi John,
Ready by end-April? That counts as a miracle! I'll let David know.
Janet

------ Original Message ------
From: "Jankowski, John" <John.Jankowski@ssa.gov>
To: "Janet Stein" <jbstein@nber.org>
Sent: 3/16/2020 1:16:54 PM
Subject: RE: FW: Re[2]: [EXTERNAL] Question about access to data from 2018 DI ASR

> Hi Janet,
>
> I heard back from our Baltimore staff on this. They can have one of our contractors put together the data that David is requesting, but it likely won't be ready until end-April. Let me know if this timeframe will work for him.
>
> Thanks,
>
> John
>
> From: Janet Stein <jbstein@nber.org>
> Sent: Tuesday, March 10, 2020 12:05 PM
> To: Jankowski, John <John.Jankowski@ssa.gov>
> Subject: Re: FW: Re[2]: [EXTERNAL] Question about access to data from 2018 DI ASR
>
> Thanks, John
>
> Here's what I heard back from David Molitor:
>
> Thanks for the update. Limiting the years to 2008 – 2017 is fine. The essence of what we're doing is comparing how things change as you cross grid rule age thresholds, so having something more granular like year of age near these thresholds is critical, hence our request for year of age between 45 and 64.
>
> Is it possible to get more clarity on the disclosure rule for suppressing cell values (e.g., if value is < 11)? It is possible that simply knowing that the value is less than this threshold will be sufficient for our needs.
>
> In the event that cell suppression under year-of-age splits creates a serious problem for our analysis, one solution might be to aggregate data by calendar year rather than by age. A downside here is that this would prevent us from analyzing how grid rule patterns change over the business cycle, but at least we would be able to conduct our primary analysis.

------ Original Message ------

From: "Jankowski, John" <John.Jankowski@ssa.gov>

To: "Janet Stein" <jbstein@nber.org>

Sent: 3/10/2020 2:51:22 PM

Subject: FW: Re[2]: [EXTERNAL] Question about access to data from 2018 DI ASR

Hi Janet,

Just following up on this one. Let me know if you want me to pursue this request.

Thanks,

John

From: Jankowski, John
Sent: Wednesday, March 04, 2020 8:39 AM
To: 'Janet Stein' <jbstein@nber.org>
Subject: RE: Re[2]: [EXTERNAL] Question about access to data from 2018 DI ASR

Hi Janet,

[Quoted text hidden]
Subject: Re[2]: [EXTERNAL] Question about access to data from 2018 DI ASR

Hi John,

Here's the reply:

We would like splits by year of age. If it’s easier or important for SSA to bin some ages, the following bins would suffice:

Ages <20,

Ages 20-24,

Ages 25-29,

Ages 30-34,

Ages 35-39,

Ages 40-44,

year of age starting at 45 (i.e., 45, 46, ..., 64)

------ Original Message ------

From: "Jankowski, John" <John.Jankowski@ssa.gov>

To: "Janet Stein" <jbstein@nber.org>

Cc: "Murphy, John" <John.Murphy@ssa.gov>

Sent: 2/28/2020 11:26:56 AM

> Subject: RE: [EXTERNAL] Question about access to data from 2018 DI ASR
>
> > Hi Janet,
> >
> > [Quoted text hidden]

David Molitor <dmolitor@illinois.edu> Thu, May 7, 2020 at 3:44 PM
To: Janet Stein <jbstein@nber.org>

Strange. No attachment is showing up.

I created a file request on Dropbox--are you able to share the file that way? Here is the link:

https://www.dropbox.com/request/ZR1hd4rZcy3Jb33rfCDH

Thanks,
David

[Quoted text hidden]

Janet Stein <jbstein@nber.org> Thu, May 7, 2020 at 3:46 PM
Reply-To: Janet Stein <jbstein@nber.org>
To: David Molitor <dmolitor@illinois.edu>

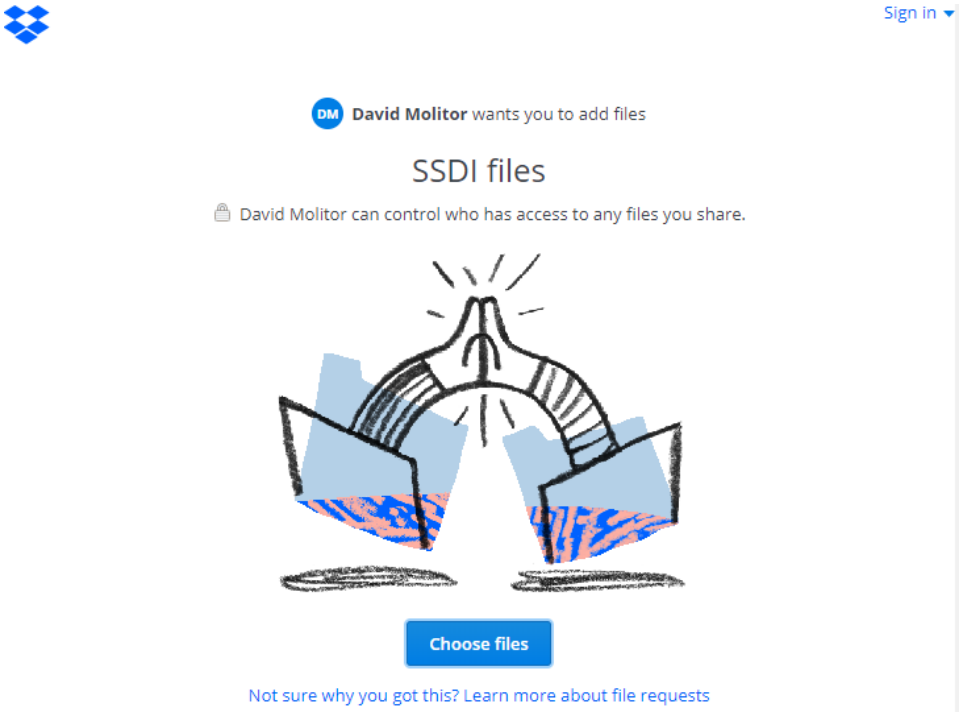
How do I get from this link to your box?

------ Original Message ------
From: "David Molitor" <dmolitor@illinois.edu>
To: "Janet Stein" <jbstein@nber.org>
Sent: 5/7/2020 4:44:54 PM
Subject: Re: [Attachment Removed] Fw: RE: Re[2]: FW: Re[2]: [EXTERNAL] Question about access to data from 2018 DI ASR

> Strange. No attachment is showing up.
>
> I created a file request on Dropbox--are you able to share the file that way? Here is the link:
>
> https://www.dropbox.com/request/ZR1hd4rZcy3Jb33rfCDH
>
> Thanks,
> David
> <http://www.davidmolitor.com>
> [Quoted text hidden]

David Molitor <dmolitor@illinois.edu> Thu, May 7, 2020 at 3:49 PM
To: Janet Stein <jbstein@nber.org>

If you click on the link, it should bring you to a landing page (screenshot below). If you click on "Choose files", you can select a file from your computer and it will put a copy in my Dropbox folder. Does this answer your question?



Thanks,
David
[Quoted text hidden]

Janet Stein <jbstein@nber.org> Thu, May 7, 2020 at 3:52 PM
Reply-To: Janet Stein <jbstein@nber.org>
To: David Molitor <dmolitor@illinois.edu>

OK. Should be all set.

------ Original Message ------
From: "David Molitor" <dmolitor@illinois.edu>
To: "Janet Stein" <jbstein@nber.org>
Sent: 5/7/2020 4:49:44 PM
Subject: Re: Re[2]: [Attachment Removed] Fw: RE: Re[2]: FW: Re[2]: [EXTERNAL] Question about access to data from 2018 DI ASR

> If you click on the link, it should bring you to a landing page (screenshot below). If you click on "Choose files", you can select a file from your computer and it will put a copy in my Dropbox folder. Does this answer your question?
>
> [Quoted text hidden]

David Molitor <dmolitor@illinois.edu> Thu, May 7, 2020 at 3:55 PM
To: Janet Stein <jbstein@nber.org>

Got it, thanks!

David
[Quoted text hidden]

Colleen Carey <colleen.carey@cornell.edu> Fri, May 8, 2020 at 10:21 AM
To: David Molitor <dmolitor@illinois.edu>
Cc: "Miller, Nolan H" <nmiller@illinois.edu>

David, did you get two zip files? I only got 1. The one I got looks like pretty much what we wanted (the tables we mentioned for single year of age, pooled across the years), although we may want to enlist someone to pull them into Excel/ Stata before David even starts to take a look at them because they are, oddly, HTML files.
[Quoted text hidden]

David Molitor <dmolitor@illinois.edu> Fri, May 8, 2020 at 10:24 AM
To: Colleen Carey <colleen.carey@cornell.edu>
Cc: "Miller, Nolan H" <nmiller@illinois.edu>

Yeah, when we asked to get the same tables, just by age, they did exactly that! I think the best thing is to develop a code that scrapes these files into a dataframe. I'm going to ask my RA to work on this, although my RA stops work in a week so I'm not sure if they'll be able to get it done before then.

Also, yes, I only got one file at first, but I followed up with Janet and got the second one as well.

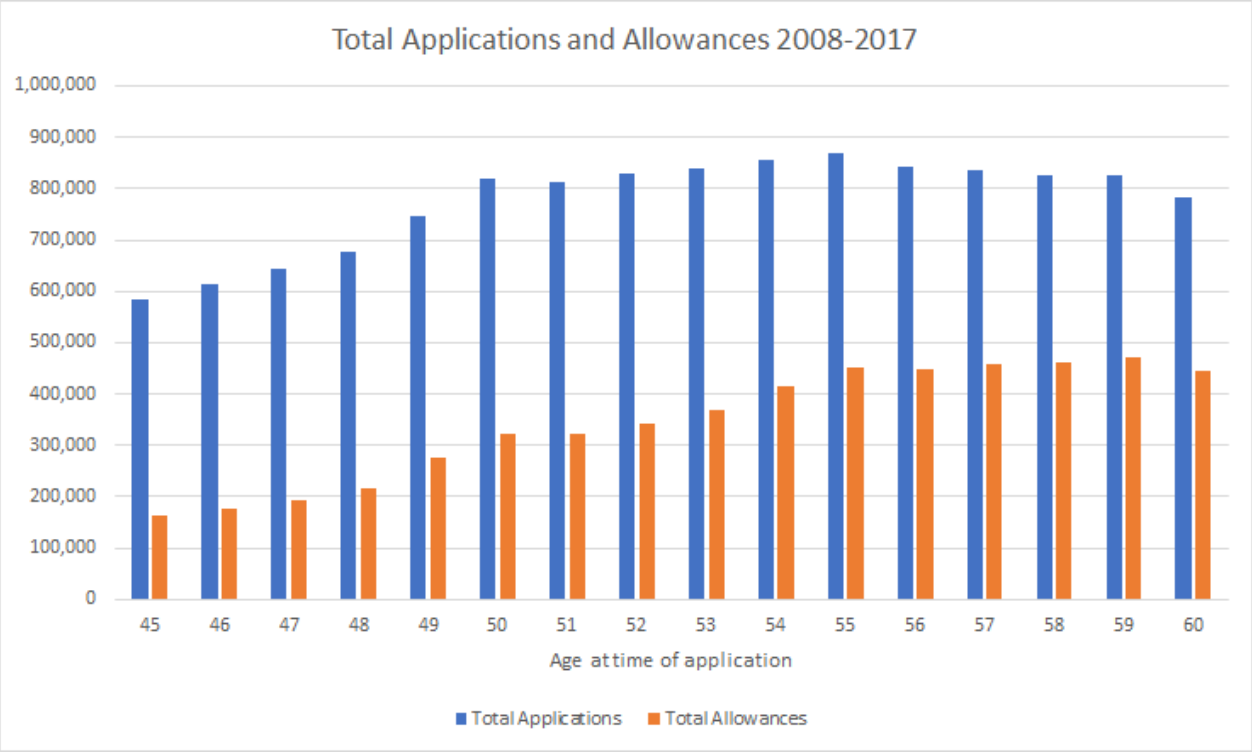
David
[Quoted text hidden]

Colleen Carey <colleen.carey@cornell.edu> Fri, May 8, 2020 at 10:27 AM
To: David Molitor <dmolitor@illinois.edu>
Cc: "Miller, Nolan H" <nmiller@illinois.edu>

Cool. It does copy-paste into nice columns in Excel (or at least it just did for me) so brute force is an option if a scraper is challenging.
[Quoted text hidden]

David Molitor <dmolitor@illinois.edu> Fri, May 8, 2020 at 11:36 AM
To: Colleen Carey <colleen.carey@cornell.edu>
Cc: "Miller, Nolan H" <nmiller@illinois.edu>

I did a quick grab of the data from ASR Table 60 for ages 45-57. I then summed up total applications and total allowances, by age, for the "All disabled beneficiaries" group. Here's what I get (spreadsheet attached):



I don't see any sharp discontinuities, but applications rise going into the age cutoffs, then drop off just after. By contrast, allowances rise almost monotonically until they dip at age 60.

David
[Quoted text hidden]

Total_Applications_Allowances.xlsx
20K

Colleen Carey <colleen.carey@cornell.edu> Mon, May 18, 2020 at 9:24 AM
To: David Molitor <dmolitor@illinois.edu>
Cc: "Miller, Nolan H" <nmiller@illinois.edu>

David, did you pull everything into Excel? Can you put it in the Dropbox? I'd like to explore a little.

Should we put another call on the calendar in hopes of maintaining momentum? (My nanny had to drop to 5h a day, but I am always free 8-1 ET and will usually be working again 4-6 ET.)

Best,
Colleen

[Quoted text hidden]

David Molitor <dmolitor@illinois.edu> Wed, May 20, 2020 at 5:51 PM
To: Colleen Carey <colleen.carey@cornell.edu>
Cc: "Miller, Nolan H" <nmiller@illinois.edu>

Hi Colleen,

I ended up writing code that scrapes the data from the tables and puts it into a rectangular format (csv). First thing tomorrow morning, I will make a do file that combines these in Stata format. At that point, it would be very helpful if you could take a look.

I'll be in touch with more details tomorrow morning.

David
[Quoted text hidden]

David Molitor <dmolitor@illinois.edu> Thu, May 21, 2020 at 12:17 PM
To: Colleen Carey <colleen.carey@cornell.edu>
Cc: "Miller, Nolan H" <nmiller@illinois.edu>

Hi Colleen,

I have finished cleaning the SSDI ASR tables. The folder containing raw and cleaned data and associated scripts can be found on Dropbox at

"SSDI/data/SSDI_ASR"

Some notes about the data:

/html - folder containing original, raw HTML files from SSA.
/csv - csv versions of each of the original HTML files (conversion performed by R scripts).
/dta - Stata dta versions of the csv files (created by the Stata script).
/final - Stata files, one per file, with cleaned, labeled variables (created by the Stata script); variables "group age year" uniquely identify observations; variables are labeled with the original column labels from the HTML files.

You should only need to pull from the /final folder. Please let me know if anything seems off or is unclear!

David
[Quoted text hidden]